DESIGN OF MONOLITHIC PROGRAMMABLE TRANSVERSAL FILTERS USING CHARGE COUPLED DEVICE TECHNOLOGY

A thesis submitted to the Faculty of Science of the University of Edinburgh, for the degree of Doctor of Philosophy

by

P B DENYER, B.Sc.

DEPARTMENT OF ELECTRICAL ENGINEERING

August 1980

ABSTRACT OF THESIS (Regulation 6.9)

Name of Candidate Peter Brian Denyer

Address 19 Craiglockhart Road North, Edinburgh EH14 1BR

Degree PhD *Date* August 1980

Title of Thesis Design of Monolithic Programmable Transversal Filters Using Charge Coupled Device Technology

The transversal filter is a widely used signal processing structure having application in both matched- and frequency- filtering systems. Monolithic realisation of this filter is a desirable goal if system cost and size are to be reduced; indeed miniature filter forms are essential to some applications.

If such a filter can be programmed by the user, then a single device might find extensive commercial application. In addition, real time programmability offers possibilities in intelligent filtering, where a system might alter its characteristics in response to a changing environment.

A high density monolithic programmable transversal filter is developed in this Thesis. The primary application of the device is as a flexible matched filter for compact, low power Sonar equipment, but its potential extends to many applications in speech, data transmission, and ultimately video systems.

A direct parallel implementation of the classical transversal filter is choosen as offering the best combination of packing density and bandwidth. The reported realisations are based on charge coupled device technology, combined with advanced linear MOS circuit techniques which together enable the complete filtering function to be formed monolithically. Among the circuit techniques developed in detail are floating-gate tapping of the CCD, linear analogue multiplication with the MOS transistor, and control and stabilisation of the CCD transfer function on chip, using a novel operational amplifier.

Two practical monolithic filters, of 64 and 256 points respectively, have been developed. They demonstrate the success of these techniques for compact matched filtering applications with Time-Bandwidth up to 1000. Signal bandwidths to 1MHz, and power dissipation as low as 1mW per point are reported.

Use this side only

DECLARATION OF ORIGINALITY

This Thesis, composed entirely by myself, reports on work conducted by myself in the Wolfson Microelectronics Institute, University of Edinburgh.

Signed

P. B. Denyer

ACKNOWLEDGEMENTS

I would like to express my sincere gratitude to Dr J Mavor and Dr A D Milne for their continued supervision and encouragement. For permission to undertake this work, and for support I gratefully acknowledge the Wolfson Microelectronics Institute, University of Edinburgh, and the device sponsors Ministry of Defence, Procurement Executive.

Sincere thanks are due to many fellow members of staff for their invaluable assistance, in particular to Dr J W Arthur and J R Reid for systems work, and to V Browne and J Gooding of Plessey (Research) Caswell Limited for their attentive device processing and design advice, which often exceeded any commercial requirement.

LIST OF CONTENTS

LIST OF CONTENTS *(continued)* *page no*

LIST OF CONTENTS *(continued)* *page no*

LIST OF CONTENTS *(continued)*

LIST OF SYMBOLS

a_n	multiplier gain variation coefficient
A	signal amplitude, CCD stage area
A_0	surface area covered by fat-zero charge packet
A_V	voltage gain function
b_n	signal offset coefficient
B	filter bandwidth, CCD voltage gain
c_n	tap gain variation coefficient
C_{A-P}	aluminium-polysilicon capacitance per unit area
C_B	blocking capacitor
C_D	total depletion region capacitance under CCD gate
C_{gb}	small-signal MOST gate-bulk capacitance
C_{gd}	small-signal MOST gate-drain capacitance
C_{gs}	small-signal MOST gate-source capacitance
C_L	total CCD tap load capacitance
C_{n+}	diffusion-substrate capacitance per unit length and area
C_0	total oxide capacitance under CCD gate
C_{ox}	oxide capacitance per unit area
C_{P-P}	polysilicon-polysilicon capacitance per unit area
d_n	signal distortion polynomial
e_n	reference offset coefficient
f	frequency (Hz)
f_c	clock frequency
f_s	sample frequency
f_n	reference gain variation coefficient
F(z)	Z transform function

g_{ds}	small signal MOST channel conductance
g_m	small signal MOST transconductance
g_n	reference distortion polynomial
h_j	tap weight coefficient
i_n	convolution error polynomial
I_{DS}	MOST drain-source current
I_0	MOST multiplier current correction
I_P	MOST multiplier product current
I_S	MOST multiplier current before correction
$\underline{k}$	convolution error vector
l_n	convolution error polynomial
L	MOST channel length
m_n	convolution error coefficient
n	integer index, tap number
N	filter length
N_A	substrate doping density
N_{SS}	surface state density
o(t)	filter output waveform
p	number of CCD clock phases
q	electron charge
Q_{dark}	charge generated by dark current
Q_S	signal charge
r	impulse residual index
r_o	small signal output resistance
t	time
T	sample period, absolute temperature, tap transfer function
T_d	total chirp duration

V_A	voltage constant
V_{DS}	MOST drain-source voltage
V_E	effective MOST gate drive ($V_{GS} - V_T$)
V_{EB}	bipolar emitter-base knee voltage
V_{FB}	flat band voltage
V_G	MOS gate voltage
V_{GS}	MOST gate-source voltage
V_T	MOST threshold voltage
V_{tap}	tap output signal voltage
V_X	multiplier voltage
V_y	multiplicand voltage
W	MOST channel width
x(t)	signal waveform
x_i	signal sample x(iT)
β, β_o	MOST transconductance factor
γ	MOST body effect constant
Δ	small change
ε	charge transfer inefficiency
ε_o	permittivity of free space
ε_s	relative permittivity of silicon
θ	MOST field-dependent mobility constant
μ_o	mobility
μ	chirp sweep frequency
ν	noise voltage
ϕ_F	Fermi potential
ϕ_i	clock phase, i = 1, 2, ...
ϕ_S	surface potential
ω	radian frequency
ω_o	chirp start frequency

To my wife Fiona.

CHAPTER ONE : INTRODUCTION

In 1969 Boyle and Smith[1] conceived the charge coupled device (CCD) as an alternative to magnetic bubble memory, for metal-oxide-semiconductor technology. The potential of this device to process analogue signals was quickly realised, and applications in signal processing were vigorously developed. Furthermore, since the charge packets could be generated optically as well as electrically, the CCD found immediate and lasting acclaim as a solid state imaging element, for both visible and infrared light.

In its simplest form, shown in Figure 1.1, the CCD comprises a row of very closely spaced MOS capacitors, between which discrete packets of charge are transferred within the semiconductor, under clock control. A fast, dense shift register function results. For signal processing, the strength of the CCD lay in this ability to realise the clock controlled delay of signals extremely compactly so that, when combined with linear multiplications and summation, many filtering functions may be realised monolithically, and at exceptionally low power levels. Savings of several orders of magnitude in size and power might be gained over conventional digital realisations. Both recursive and transversal CCD based filters are feasible, although the former appear more sensitive to imperfections in the CCD, and in particular to any parametric drift. It is thus the transversal filter realisations that have most successfully exploited the useful properties of CCD.

An excellent example of the application of CCD to signal processing is the transversal split-gate filter[54], shown in Figure 1.2,which

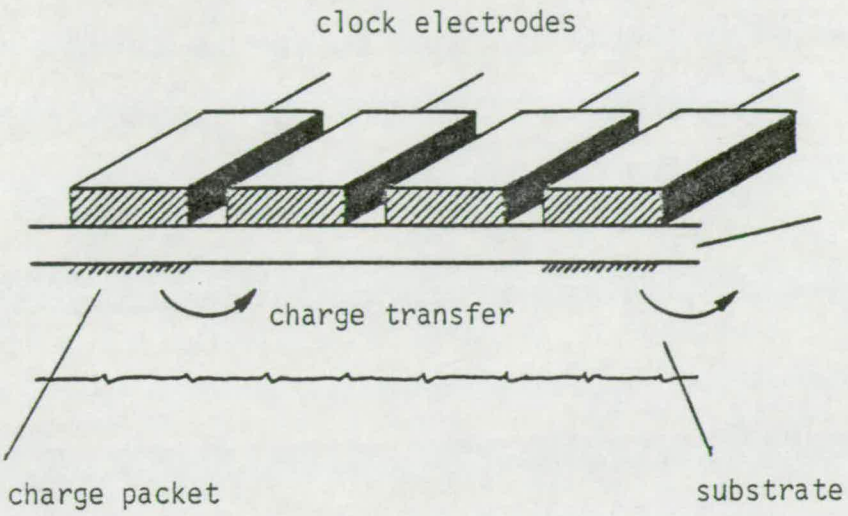


Figure 1.1: The Charge Coupled Device

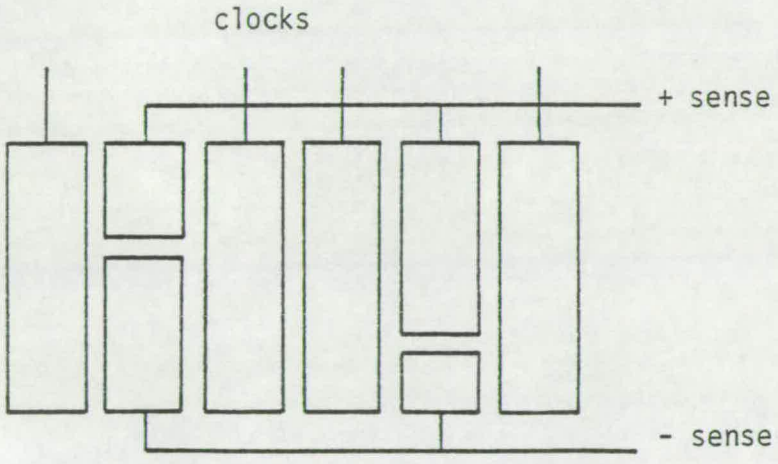


Figure 1.2: Split-gate CCD Transversal Filter

combines all of the delay, multiply and accumulate functions in a single structure. Here multiplication is attained by detecting some fraction of a charge packet, through a split sensing capacitor within the CCD structure. Accumulation occurs through a common connection of these sensing nodes, whilst the CCD provides the delay of charge packets between the sensing sites. A disadvantage of the split-gate filter is that it has a fixed characteristic, which may only be altered at the mask-making stage. Thus a high fixed cost pertains to any individual application, and the filter characteristics may not be altered *in situ*, except by multiplexing devices. Now if such a filter could be electrically programmed in real time, several new applications would be feasible.

Firstly, a single device could be field-programmed to perform a variety of tasks without the expense and delay of refabrication. Secondly, such a device might be controlled to alter its characteristic in response to a changing environment, especially in a mobile application. An extension to these applications is to be found if the device is continuously operating in a feedback loop to intelligently alter its characteristic, to produce a programmed response from an arbitarary input[96].

The objective of this research is the development of a practical programmable filter using the compact delay-line advantage of CCD, combined with advanced linear MOS circuit techniques to realise the programmable multiplication and accumulation functions. This topic has been a subject of research interest at Edinburgh since 1973. In 1976 MacLennan published his Thesis[11] covering much of the early development work on non-destructive CCD tapping and MOS multiplication techniques. At that time, however, no monolithic realisation had

been demonstrated. These early techniques are refined here, and new methods are developed that enable a complete integration of the filter.

The development programme culminated in 1977 with the demonstration of a monolithic programmable CCD filter[59] (PTF I) containing 64 sampled data points. Much experience was gained from this realisation, which lead to the development in 1979 of an advanced 256 point filter[65,71] in an improved CCD technology. It is the major aim of this thesis to report on the design and characterisation of these devices, and to compare them with contemporary realisations.

The remainder of this introduction is devoted to a summary of the transversal filter, and to its application as a sonar matched filter, for which the devices reported in this thesis have been primarily developed. Chapter 2 contains a resume of CCD-based filter architectures, from which the particular structure best suited to the sonar application evolves. Chapter 3 covers the basic CCD and linear MOS theories essential to these circuits. In Chapter 4, the tapped CCD delay line is examined in detail and illustrated with two practical examples, and the concept of feedback linearisation is introduced as a means of idealising the CCD transfer function. Chapter 5 deals in depth with an essential filtering function, linear multiplication, using the MOS transistor; a novel single transistor multiplier of superior performance is presented. An operational amplifier suitable for integration with the CCD, and in particular for use in feedback linearisation, is developed in Chapter 6. Chapters 7 and 8 cover the detailed design and characterisation of the two device realisations, whilst Chapter 9 considers some of the limitations of the structures in the light of their results, with particular attention to the noise performance.

1.1 THE TRANSVERSAL FILTER

The transversal filter is a finite impulse response structure, shown in its most common, or *direct*, form in Figure 1.3. Input samples x(nT) are delayed, and the outputs of the successive delays are weighted and summed to form the filter output. The impulse response of the structure is simply the series of tap weighting coefficients, $h_1, \ldots, h_N$. The response of the filter to a sampled input sequence, x(nT), is given by its convolution with $\underline{h}$;

$$o(nT) = \sum_{j=1}^{N} h_j x(nT - j), \qquad (1.1)$$

where N is the number of points in the filter.

Inasmuch as the h_j are to be programmable, the filter may also be considered as a correlator, of $\underline{x}$ with $\underline{h}$. This structure finds application in two main areas; as a matched filter or correlation detector, and as a frequency filter.

1.2 MATCHED FILTERING

The matched filtering theorem[88] states that in order to detect a signal in the presence of white noise with the optimum signal-to-noise ratio, a filter having an impulse response equal to the time inverse of the signal should be used. This gives the tap weights a spatial distribution along the filter which is the same as the shape of the signal along the time axis. At some time the signal 'matches' the tap profile, and the various weighted outputs add coherently, whilst the noise always adds incoherently.

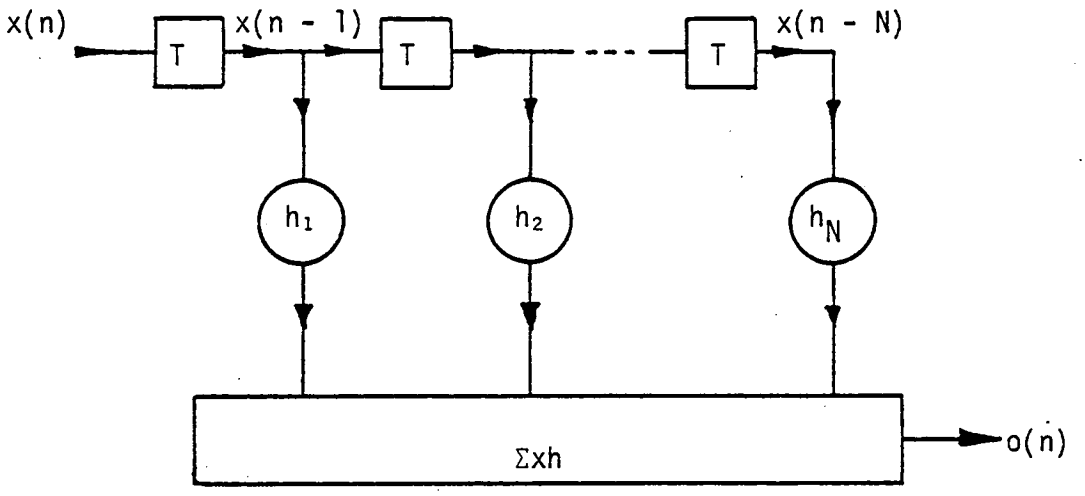


Figure 1.3: The transversal filter

At the instant of correlation the normalised peak output signal is

$$o_p = N\overline{h_n^2} , \tag{1.2}$$

whereas the normalised r.m.s. noise level is

$$o_n = \sqrt{N\overline{h_n^2}} . \tag{1.3}$$

The SNR at the output has therefore improved by a factor $\sqrt{N\overline{h_n^2}}$, or $\sqrt{N/2}$ for sinusoidal signals. This improvement is known as the 'processing gain' of the filter.

1.3 CHIRPS

The 'chirp' waveform is commonly used in matched filter systems because it combines a good time compression characteristic with a high tolerance to frequency distortion (Doppler Shift). A typical chirp, and its matched response are shown in Figure 1.4. Note how the energy of the chirp is effectively compressed into a narrow peak; it is this peak that is more easily detected in the presence of noise.

Chirps have the form;

$$[f(t)]_{0<t<T_d} = A \cos(\omega_o t \pm \mu t^2 + \emptyset) , \tag{1.4}$$

where A is the constant amplitude, and the $\pm$ sign distinguishes 'up' and 'down' chirps. The chirp sweeps linearly over a frequency range

(a) chirp

(b) autocorrelation function

Figure 1.4: The chirp waveform and its autocorrelation function

of μ/π in time T_d. Now the autocorrelation function may be approximated (over its central portion) by a carrier,

$$A \cos(\omega_o t \pm \mu(\frac{T_d}{2})t + \emptyset), \tag{1.5}$$

modulated by a sinc x function of the form

$$\frac{\sin \mu t}{\mu t} \tag{1.6}$$

The width of the main correlation peak is inversely proportional to the swept bandwidth $B = \mu/\pi$. The sinc x function exhibits sidelobes at an initial significance of -13dB. In some applications these can mask a weaker adjacent signal, so the tap weights may be shaped by a weighting function to reduce the sidelobe significance, at the expense of a marginal increase in peak width. In any case, for maximum resolution of the returns in time at the detector, the swept bandwidth should always be as high as possible, but never above $f_c/2$ for a sampled data filter.

Now the processing gain is determined by the length of the filter,

$$N = f_c T_d, \tag{1.7}$$

so for a system sampling at the Nyquist rate, $f_c = 2B$;

$$N = 2\ BT_d. \tag{1.8}$$

This gives a direct link between the filter length N and the maximum time-bandwidth product that may be processed. This product is often taken as a measure of the performance of the system, combining as it does the processing gain and resolution capability. Evidently large values of N are desirable.

1.4 CHIRP DETECTION SYSTEMS

Because chirps are often transmitted on a carrier, the phase term, $\emptyset$, is normally indeterminate and can destructively effect the correlation peak, especially if the carrier frquency is of the same order as the swept bandwidth. Various system configurations which cancel this effect are known[83], generally involving quadrature channels with mixing to remove the carrier term. An example is shown in Figure 1.5. It is not the purpose here to examine these systems in any detail however, but rather to consider the matched filter element in isolation. Thus, the devices are often tested as single channel filters responding to correctly synchronised chirps, with no phase indeterminacy.

1.5 THE SONAR PROBLEM

Active sonar is used to detect underwater features by sound transmission and reflection. A sound signal is transmitted into the water, and detected after reflection from a target. The total time delay is measured to determine the range of the target. In fact, because the target may be relatively large (compared to the radar situation), a distinct group of returns is often seen, which may be used to identify the target to some extent. As with radar, matched filtering techniques are used to extend the sonar range by detecting spread-energy signals buried in noise. In fact the sonar problem is more complicated than this. Temperature gradients in the sea, surface waves and sea-bed features all give rise to excessive (time-dependent) clutter, or reverberation, including multiple reflections, and there is a further time-variance in the transmission characteristic which results in coherent multiplicative noise. For these reasons,'shaped' filters with reduced sidelobe features are not

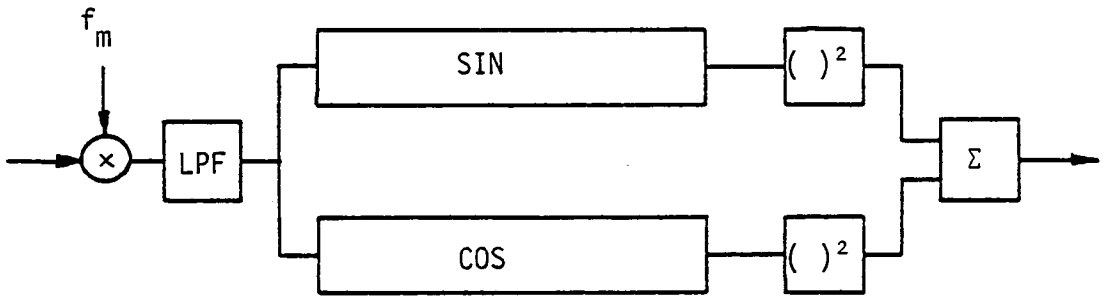


Figure 1.5: Quadrature channel chirp detection system

particularly useful, and emphasis is given to maximum range resolution, which is obtained with unweighted responses. Sampling rates up to 50 KHz, and BT_d up to 1000 are desirable operating parameters for a usefully flexible system.

Where the sonar system is ship-borne digital hardware is feasible, offering arbitrary accuracy at the expense of size, weight and power dissipation. A light-weight, possibly mobile, independent equipment may not be able to support the digital circuitry, as size and power dissipation can be at a premium. It is for these applications that the CCD-based programmable filters reported in this thesis have been developed. The requirement is evidently for a maximally compact device combining as many filter points as possible at low power levels. These considerations are paramount in determining the choice of filter architecture and technology.

1.6 FREQUENCY FILTERING

Matched filter applications notwithstanding, a truly programmable filter may also be used for frequency filtering. The tap weights are programmed to form as an impulse response the inverse Fourier transform of the desired frequency characteristic. The transversal filter is particularly useful for forming all-zero, linear-phase responses[74], but may be programmed to realise an arbitrary response, given sufficient points and tap weight accuracy.

1.7 TECHNOLOGY

It is important to realise that any practical device realisation must be based on a readily-available device technology. All of the devices in this work have been fabricated by Plessey Research (Caswell)

Limited, on their CCD processes. Initially a metal gate thin-'field' process with a single transistor threshold was available, and the first devices were developed on this. This process was discontinued in 1977 however, and replaced by a more sophisticated thick field, self-aligned, polysilicon-gate process with a range of transistor thresholds.

The later designs, in particular the 256-point filter (PTF II) were fabricated on this process. Details of the layout rules and process parameters are given in Appendix I.

CHAPTER TWO : TRANSVERSAL FILTER ARCHITECTURES

The import of the programmable transversal filter is well reflected in a wide variety of schemes which have been proposed to realise it.[45-71] No single realisation is suited to all applications, each striking a definite compromise between bandwidth, power dissipation, accuracy and size. This chapter examines the alternative architectures, and reviews some of the significant realisations which have been reported. From this knowledge, the filter architecture best suited to the compact sonar requirement is derived.

The purpose of all of these filters is some realisation of the simple convolution sum given in (1.1). Fundamentally, this requires storage of the N weighting coefficients, and of N samples of the signal waveform to be convolved. There is also the computational requirement of N multiplications and additions per filter cycle.

2.1 ARCHITECTURE

Transversal filter realisations fall broadly into categories of serial and parallel form. Serial form processors recognise the multiplier as a complex component, and minimise this hardware requirement by multiplexing a single element to perform sequentially all of the N multiplications required within one filter cycle. The bandwidth of these processors is necessarily restricted by the rate of the multiplier. Parallel processors rely on simultaneous parallel computation and accumulation of N multiplications. The filter sample rate may thus equal the multiplier rate at the expense of hardware complexity.

The classical parallel-form filter architecure, shown in Figure 2.1, is based on a serial-in-parallel-out (SIPO), or 'tapped delay line' structure. The output of each tap is coefficient-weighted and the individual products are summed. Several integrated filter realisations based on this popular architecture have been reported. The tapped delay line need not necessarily comprise one element, however, so long as the appropriate signal delay functions are attained. Figure 2.2 shows how separate, untapped delay lines might be used, thus avoiding the implementation of taps which may adversely affect the high speed performance of a single delay line.

The individual transfer functions of each of the weighted delay strings are not altered by siting the coefficient multipliers at their inputs, as in Figure 2.3. Indeed the outputs from the delay lines may be more easily summed than those of the multipliers. A commercial device using exactly this architecture has recently been reported by Weckler.[64]

In certain cases this simplified summation procedure may be incorporated at each input of a parallel-in-serial-out (PISO) structure, as shown in Figure 2.4. Such structures may be readily integrated, and the high speed programmable filter reported by White et. al[68] uses this architecture.

A clever extension of this technique to serial form using a single multiplier has been reported by Gooding[62], and is shown in Figure 2.5. The weighted input samples are here formed serially and stored in a SIPO at a rapid clock rate, before being presented in parallel to the PISO at the normal filter sample rate. Note that a recirculating (and therefore digital) coefficient memory is required.

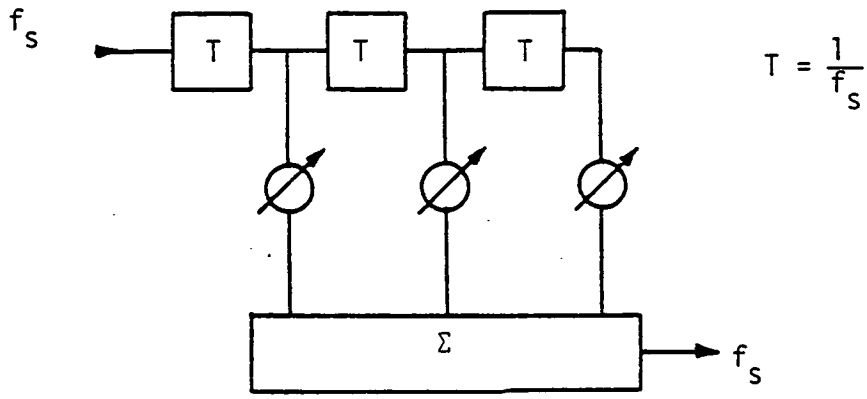


Figure 2.1: Classical parallel form architecture based on tapped delay line

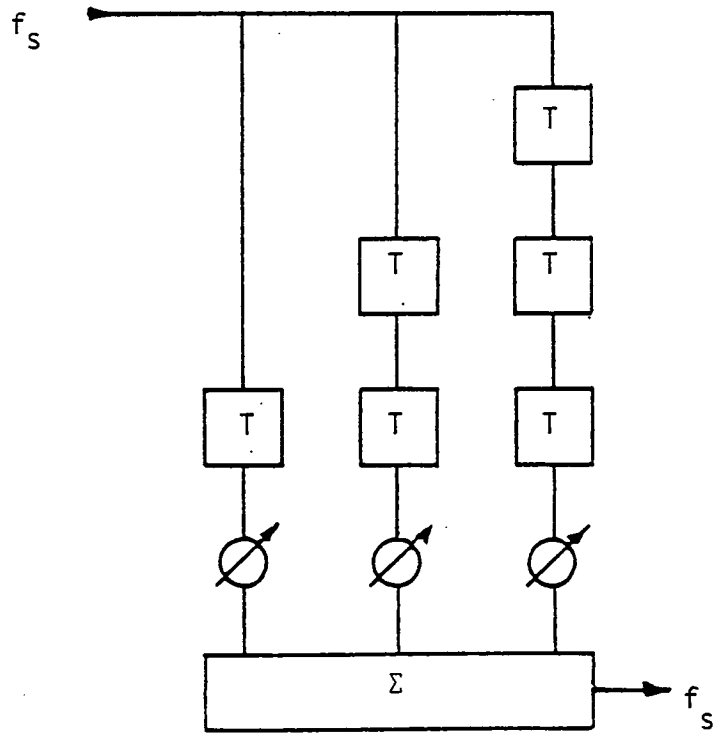


Figure 2.2: Parallel form architecture using multiple untapped delay lines

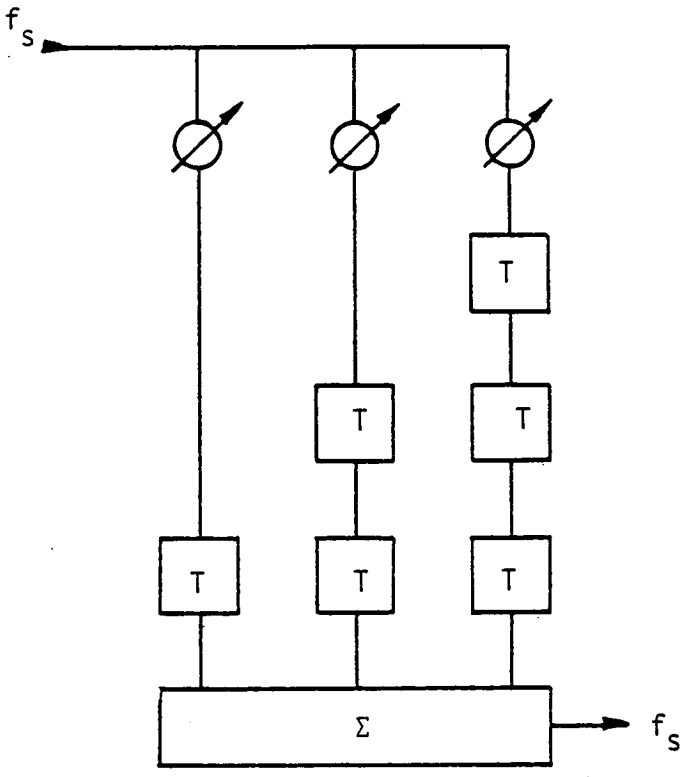


Figure 2.3: Parallel form architecture using multiple delay lines with weighted inputs

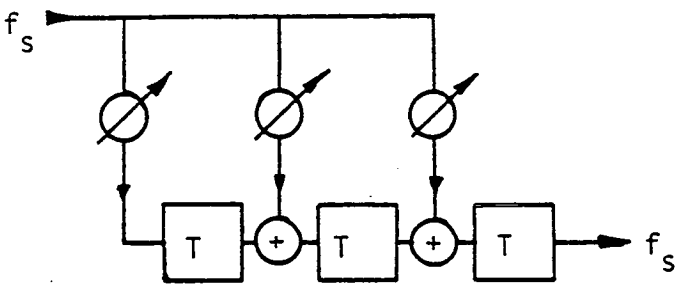
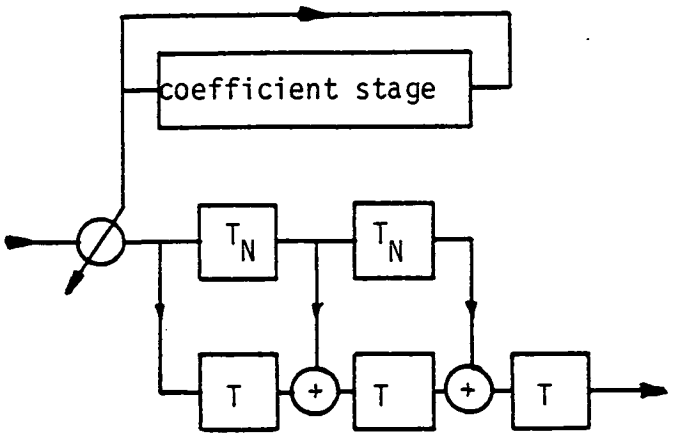


Figure 2.4: Parallel form architecture using SIPO structure



$$T_N = \frac{T}{N}$$

Figure 2.5: Serial form architecture using separately clocked SIPO and PISO structures

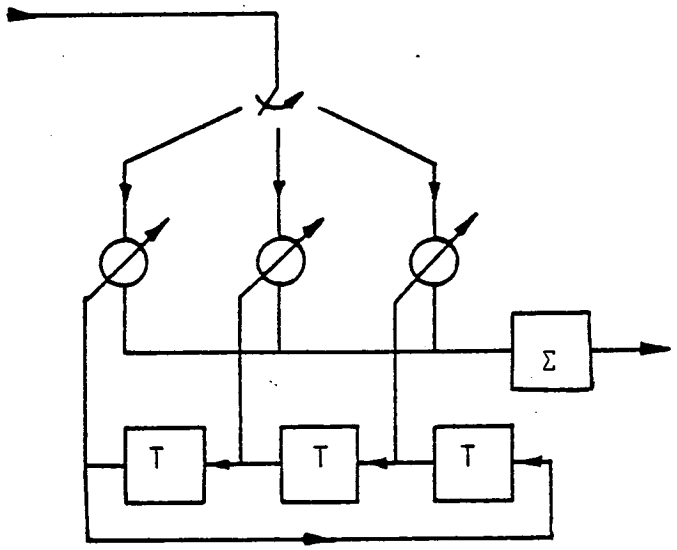


Figure 2.6: Parallel form architecture with recirculating coefficient store

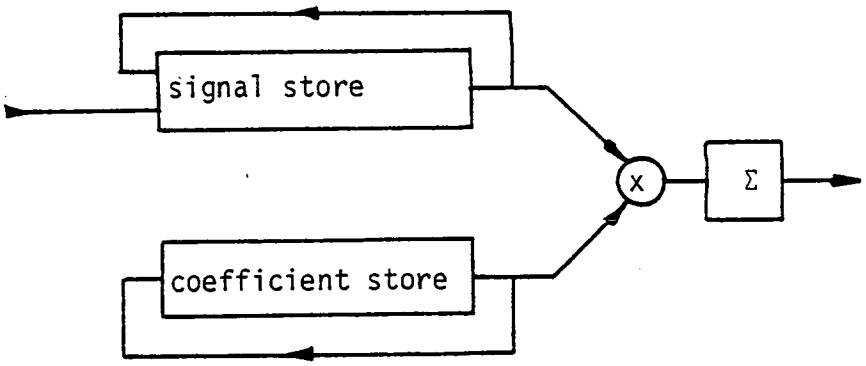


Figure 2.7: Fully recirculating serial form architecture

Figure 2.6 shows a further architecture using a recirculating coefficient store, but with parallel multiplication. This structure has been adopted by Haque and Copeland[57] for integrated filter realisations not requiring an analogue (signal) tapped delay line. This feature obviates the CCD requirement, and so may utilise a simpler linear MOS technology.

The final architecture shown in Figure 2.7 is a serial form employing recirculating signal and coefficient registers. This is a well known system configuration[56], but no integrated realisation has yet been reported.

2.2 ANALOGUE VERSUS DIGITAL PROCESSING

Convolvers based on purely digital components suffer from the disadvantage that when they are used to process analogue signals as is often required - analogue to digital (a./d.) conversion is implied, at least at the signal input. For integrated digital realisations the size and power of the multiplier restrict the choice of architecture to the serial form , and the filter bandwidth and number of points must be limited by the maximum rate of this multiplier. Digital storage however is conveniently simple and reliable, and may be extended arbitrarily to any required accuracy. However, the integrated area of these storage cells, and of the digital multiplier, are severe deterents to any monolithic realisation.

Processors using analogue circuit techniques benefit from dense, fast processing at the expense of limited accuracy and dynamic range. The performance of analogue circuitry is also less well contained than equivalent digital systems. However, more direct functional realisations are feasible and this advantage may outweigh the inherent

problems for certain applications.

A specific advantage of the use of purely analogue techniques is the feasibility of direct analogue-analogue multiplication using one or two MOS transistors[26,27,59]. The spatial economy of these elements allows the integration of dense parallel form processors with corresponding bandwidth advantages. Furthermore analogue coefficient storage is implied, and MOS capacitors offer a compact solution.

In the same technology bucket-brigade[34] and charge-coupled device[1] analogue shift registers permit the integration of SIPO and PISO elements which form the bases of many parallel processor realisations. Information within these delay lines is progressively degraded however, making recirculating analogue stores largely impractical. In addition, this degradation generally imposes an upper limit on the time-bandwidth product of signals to be processed in non-recirculating delay lines. However, values of BT_d up to 1000 are normally feasible, so only exceptionally long matched filter applications are precluded. Thermal leakage within CTDs further limits the total signal delay that may be associated with any one device, to around one second at room temperature. The faster recirculation and regeneration rates of the serial digital processors generally obviate this problem.

The best features of both analogue and digital techniques may be combined by using an analogue signal channel and a digital reference, in a parallel form realisation. The multiplier requirement is relaxed where a single bit reference is suitable and, consequently the hardware complexity of the processor is reduced. References of higher resolution may be effected by running these units in parallel.[54] Alternatively, a multibit reference may be used with multiplying d./a. converters at each filter point.[64] The complexity of this approach

Convolver type		Speed	Packing Density	Accuracy	Features
Configuration	Mode				
Serial	digital	medium-slow (trade-off with packing density)	medium-high (trade-off with accuracy)	to any requirement (with digital multiplication)	permanent signal and reference memory
	analogue and digital			limited by analogue multipliers	restricted by recirculation of analogue data
	analogue				
Parallel	digital	fast	low	to any requirement (with digital multiplication)	permanent signal and reference memory
	analogue and digital		medium	limited by analogue multipliers	permanent reference memory
	analogue		high		restricted by analogue reference decay

Table 2.1: Summary of convolver realisations

limits the filter packing density, but a good combination of bandwidth and accuracy may be achieved.

The features of these various combinations of digital and analogue, and serial and parallel realisations are summarised in Table 2.1.

2.3 A SPECIFIC ARCHITECTURE

From Table 2.1 many tradeoffs are evident; there is clearly a wide choice available to suit any particular signal processing requirement. A prime aim of this work is the integration of a filter with a maximum number of points, subject to the restrictions that the device be cascadeable, and capable of attaining multibit coefficient accuracy at sonar bandwidths. Under these conditions the parallel, analogue architectures offer the greatest packing density. Of these, architecures based on multiple delay lines clearly suffer an area disadvantage, as do PISO-based structures, especially where large numbers of filter points are to be integrated. This arises because the delay-line area must be allowed to increase along with the amount of charge that is being summed.

This not only leads to a direct area penalty (which increases in arithmetic progression) but also complicates the layout of the remainder of the chip because of the triangular shape which emerges.

Conversely, the SIPO element allows a single charge packet to be detected, or tapped, many times within the same delay line, which maintains a uniform size. Here taps may be added by simply extending the delay line, with a minimal linear increase in layout area. The potential restriction of bandwidth which was previously noted is not of concern at sonar frequencies.

The programmable filter architecture to be developed here is thus a parallel, analogue form based on a SIPO delay-line structure offering maximum integration packing density.

The chosen filter architecture shown in Figure 2.8 is both simple and compact, involving the minimum of signal manipulation and integrated silicon area. It is a *direct* realisation of the classical transversal filter and can be implemented using linear CTD and MOS component technology.

The signal register is formed from a tapped analogue CCD-delay line, which simultaneously realises the signal storage and time-shift operations required in the correlation sum. CCD realisation of this element is optimum for these applications, in that the analogue time-delay and shift processes are achieved naturally, with the most economical use of silicon area.

Because the CCD signal register provides the necessary time-shift process, a stationary analogue reference register is sufficient, supplying the weighting values to the multipliers in parallel form. An electrically simple and physically compact realisation of this element uses discrete MOS capacitors for analogue voltage memory. These feed the multiplier reference terminals via buffer amplifiers, and reference values are updated individually via a single, digitally multiplexed analogue input bus.

This static analogue reference memory implies a simpler chip structure than filters which employ both dynamic signal and reference registers. Note that it is possible to exchange the signal and reference register functions (such that the signal is loaded into the static register, and the reference coefficients are time shifted and recirculated) without changing its function.

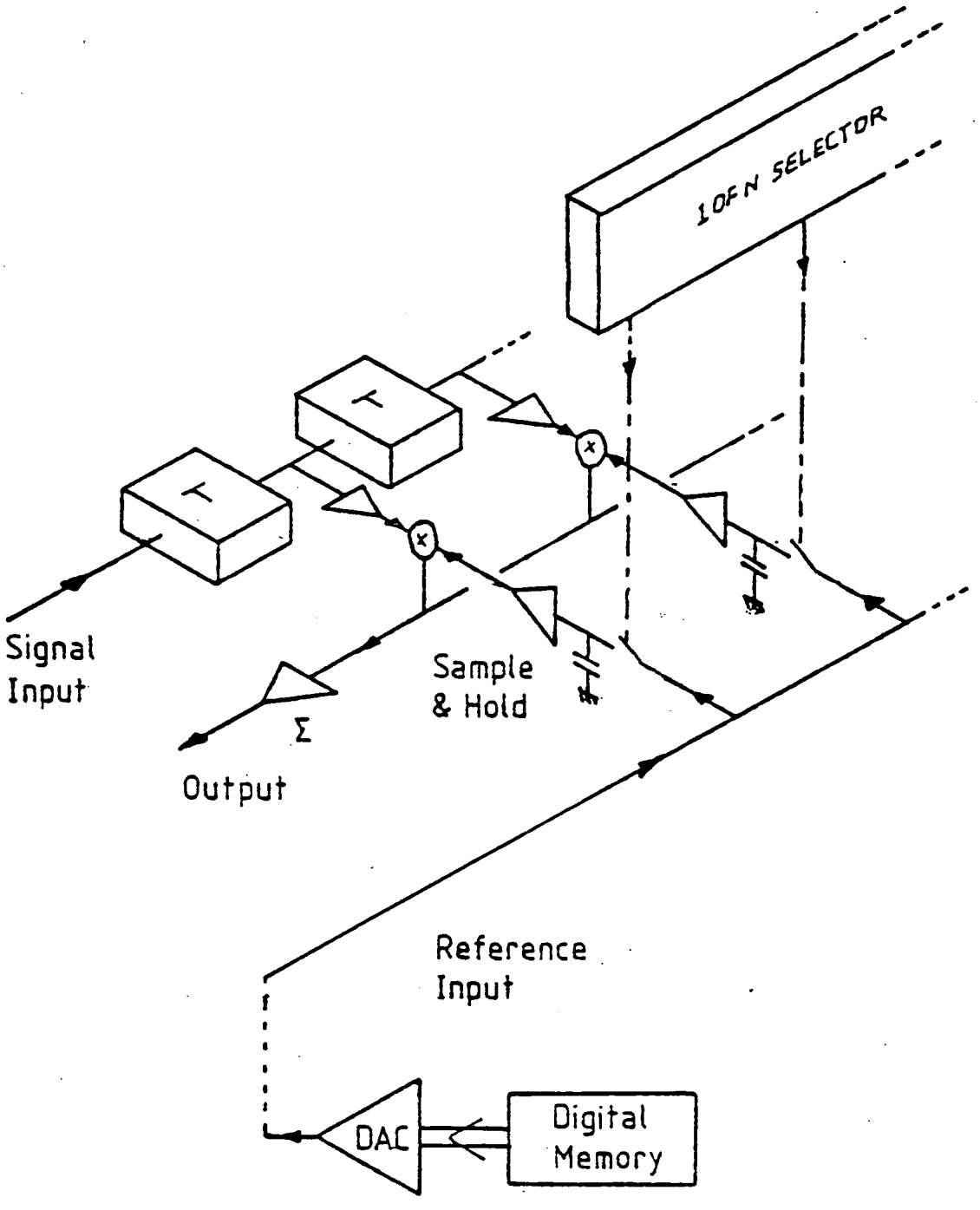


Figure 2.8: Block diagram of parallel form, analogue/analogue PTF

As previously discussed, analogue coefficient storage is by nature dynamic and thus may require some form of dedicated, external memory for refresh purposes, as shown in Figure 2.6. Despite this two-level memory requirement, the arrangement is still optimally compact for filters of more than (approximately) 32 points, because of the relative simplicity of the analogue multipliers compared with the multiplying DAC structures implied with a single digital-reference memory architecture.

For general purpose signal processing, accurate multiplication of the signal and reference (weighting coefficient) samples is required at each filter point. Now it is well known that MOS transistors, operating in the "triode" region, potentially offer a compact multiplying element. Realisations have suffered, however, from poor accuracy and dynamic range in addition to long-term stability and drift problems. For this reason, a multiplication arrangement has been developed around a *single* MOS transistor which gives much improved performance over contemporary realisations.

The operation and performance of each of these elements is described in detail in the following chapters. Before this however, a chapter on CCD and MOST device theory is included, so that their performance may be better understood.

CHAPTER THREE : CCD AND MOST FUNDAMENTALS

The potential of MOS technology to realise linear analogue circuits, together with CCD delay lines, makes it extremely suitable for LSI implementation of the programmable transversal filter. MOS technology in the UK has developed to the extend that CCD, and linear and digital MOS circuitry may be successfully combined on the same chip.[35,37] Since all of these aspects are fully exploited in a programmable filter realisation, it is important that their fundamental theory be clearly understood. This chapter sets out that theory, and proposessome simple engineering design models.

3.1 THE MOS CAPACITOR

The operation of all MOS elements stems from the theory of the MOS capacitor, shown in cross-section in Figure 3.1. The advantage of this simple Metal-Oxide-Semiconductor structure lies in the unique ability of the gate to electrically control a group, or flow, of minority carriers at the surface of the semiconductor underneath. All of the devices reported here use a p-type silicon substrate and nomenclature is consistent with this throughout. The gate material may be aluminium, or heavily doped polysilicon, and the insulator is commonly silicon dioxide.

For low gate voltages no useful operation results, but as V_G is increased, *majority* carriers within the semiconductor are repelled from the surface to form a depletion region. As the gate voltage is further increased,past a notional 'threshold' voltage,free

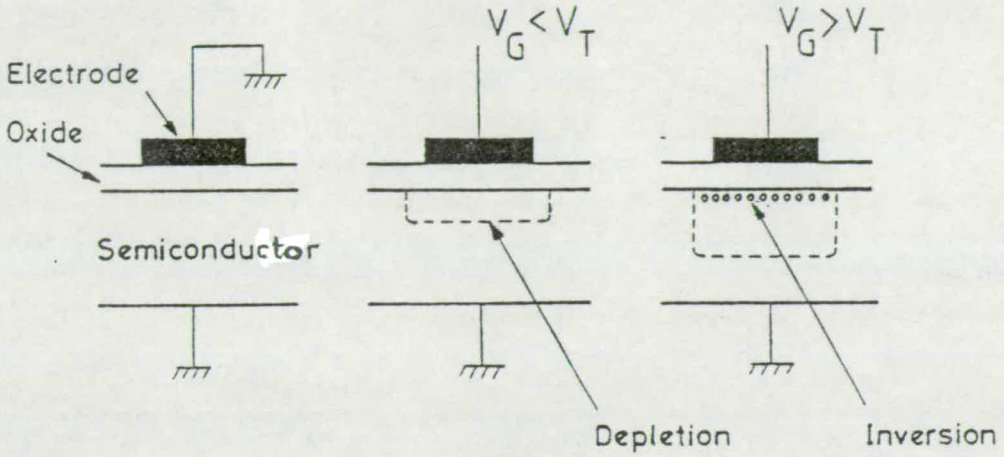


Figure 3.1: The MOS capacitor

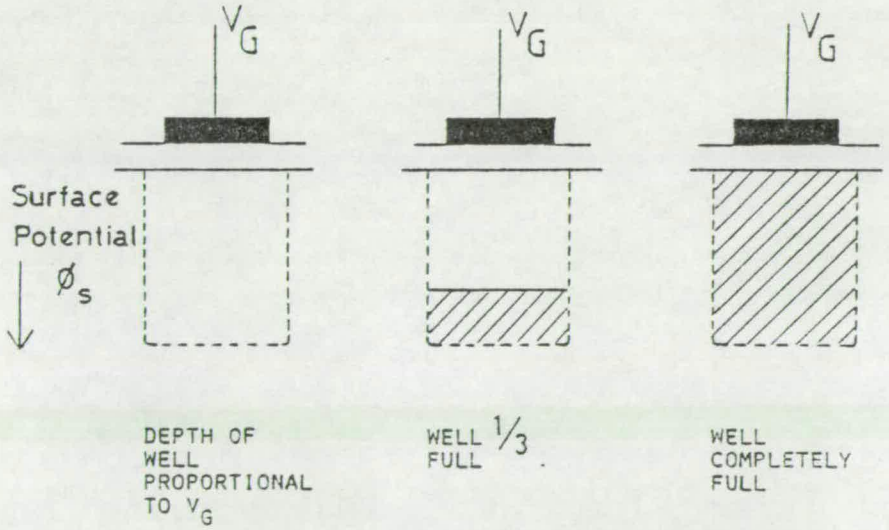


Figure 3.2: The potential well concept

minority carriers (electrons) become attracted to the electrode, and gather in an inversion layer at the semiconductor surface. In the case of the MOS transistor, these minority carriers originate from nearby n^+ diffusions. For the CCD, packets of carriers are transferred between closely spaced MOS capacitors. Note that in all cases the inversion charge layer is isolated from its surroundings by the depletion region, and is totally under the control of the gate voltage.

A mathematical relationship may be determined between the gate voltage V_G, the semiconductor surface potential ϕ_S, and the quantity of charge beneath the gate Q_S. This relationship has been well studied by Sze[2] and others. By integrating the one-dimensional Poisson equation twice about a line perpendicular to the surface, the following result may be obtained;[8]

$$\phi_S = (V_G - V_{FB}) + \frac{Q_S}{C_0} + V_A - \{2V_A [(V_G - V_{FB}) + \frac{Q_S}{C_0}] + V_A\}^{\frac{1}{2}}, \quad (3.1)$$

where V_{FB} is the so-called 'flat-band' voltage, at which inversion commences, V_A is a process dependent constant given by $N_A q \varepsilon_o \varepsilon_s / C_0$, and C_0 is the oxide capacitance beneath the gate. V_A is small (typically 0.3V), and so the square root may usually be neglected at normal working voltages, so:

$$\phi_S \simeq (V_G - V_{FB}) + \frac{Q_S}{C_0} + V_A. \quad (3.2)$$

In simple terms, the surface potential *increases* with increasing gate voltage, but *decreases* with the amount of (-ve) charge stored.

3.2 THE POTENTIAL WELL CONCEPT

The relationship of equation (3.2) leads to the concept of the 'potential well' illustrated in Figure 3.2. The surface potential $\emptyset_S$ is plotted downwards underneath the electrode, and the 'depth' of this well is initially determined by the voltage applied to the gate. As charge is trapped, the surface potential falls and the well 'fills up'. The trapped charge is thus visualised as a liquid filling the potential well created underneath the electrode.

3.3 CHARGE TRANSFER AND THE CHARGE COUPLED DEVICE

In a search for an electronic analogue to magnetic bubble devices, Boyle and Smith[1] conceived in 1969 the Charge Coupled Device (CCD). The principle of the CCD is that of a controlled manipulation of discrete packets of minority carriers between closely spaced MOS capacitors. Allowing the amplitude of these charge packets to represent sampled signal information implicates the CCD as a direct vehicle for signal processing.

Consider the directional transfer of charge packets along a series of closely spaced electrodes, as shown in Figure 3.3. We begin at time t_1, with one of the electrodes at a high potential (on), with a charge packet held in the resultant potential well. If an adjacent electrode is now turned on, t_2, the potential wells overlap and charge flows between them until the surface potential is in equilibrium. If the first electrode is now turned off, t_3, the remaining charge is all transferred to the second electrode. In order to prevent the backward flow of charge, and to ensure isolation from the following charge packet, a third electrode is required. In fact this type of electrode structure requires a minimum grouping of 3

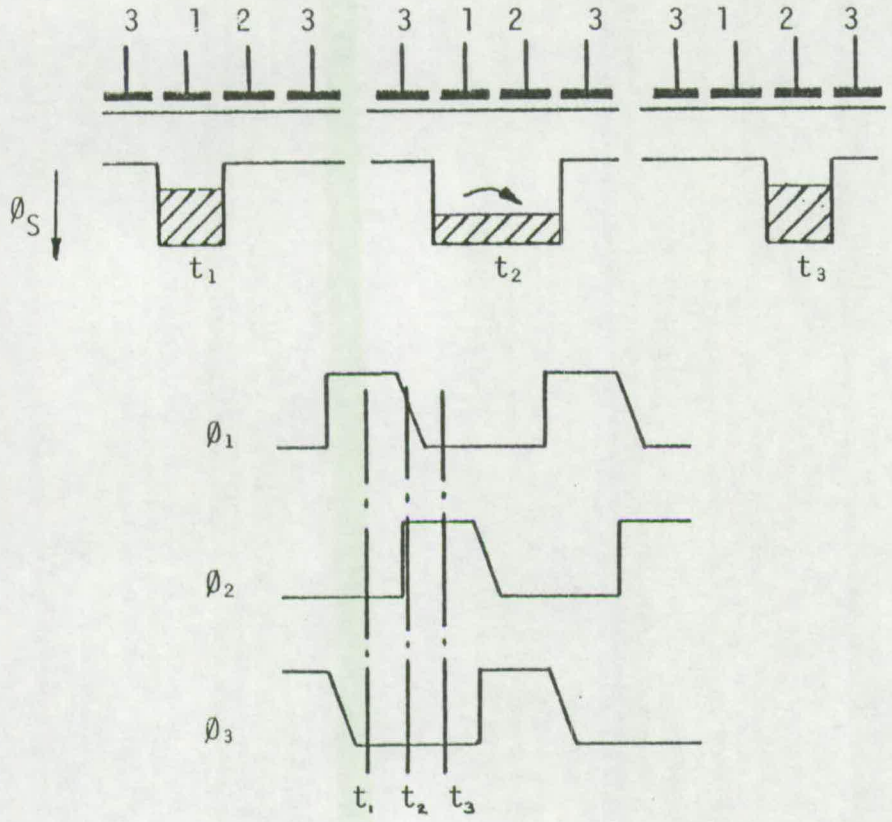


Figure 3.3: Charge transfer in a 3-phase CCD

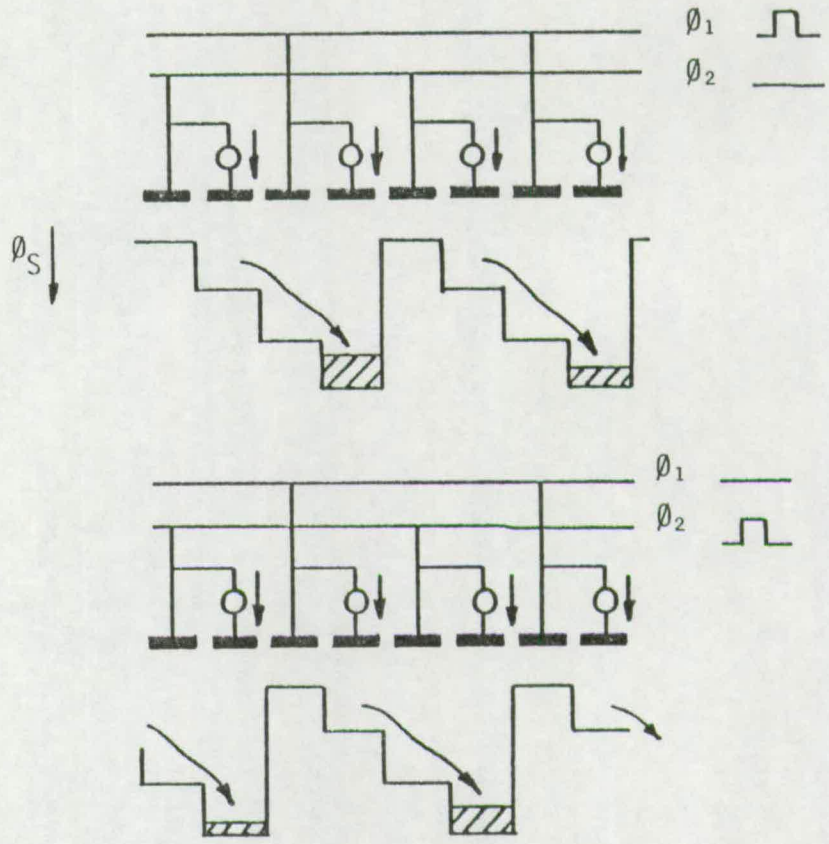


Figure 3.4: Charge transfer in a 2-phase CCD

electrodes per charge packet, commonly driven by a three-phase overlapping clock. In this way, charge is transferred between the phase 1 and phase 2 electrodes, then between phase 2 and phase 3 and so on, cyclically along the register.

The 3 phase system may be simplified by introducing asymmetrical electrodes which automatically block backward charge transfer; the arrangement is shown schematically in Figure 3.4. Each electrode now consists of two parts separated by a potential step; charge packets naturally prefer that part at the higher potential, whilst that at the lower potential blocks backward charge flow. Such a scheme requires only two non-overlapping clock phases, which are simpler to generate. The potential step is achieved in practice either by depositing the electrodes over oxide steps (remembering that $\phi_S \propto C_0^{-1}$), or by using a selective channel implant to alter the surface potential.

Both charge transfer schemes may benefit from a trick which allows a p-phase CCD to be driven by a (p-1) phase clock, again simplifying drive requirements. The technique involves biasing one phase-group of electrodes at a d.c. potential, approximately midway between the clock-on and clock-off potentials. Charge is transferred underneath these electrodes when the preceding electrodes turn *off*, and is transferred from them to the following electrodes when they turn *on*. The operation of a conventional 2 phase device in this way is shown in Figure 3.5. Clearly this represents the simplest device to operate, as only a single-phase clock waveform is required. The technique is often referred to as pseudo (p-1) phase, or as $(p-1)\frac{1}{2}$ phase clocking. E.g. the 2-phase structure becomes pseudo-1 phase, or $1\frac{1}{2}$ phase.

A further distinction is to be made, between surface- and buried-

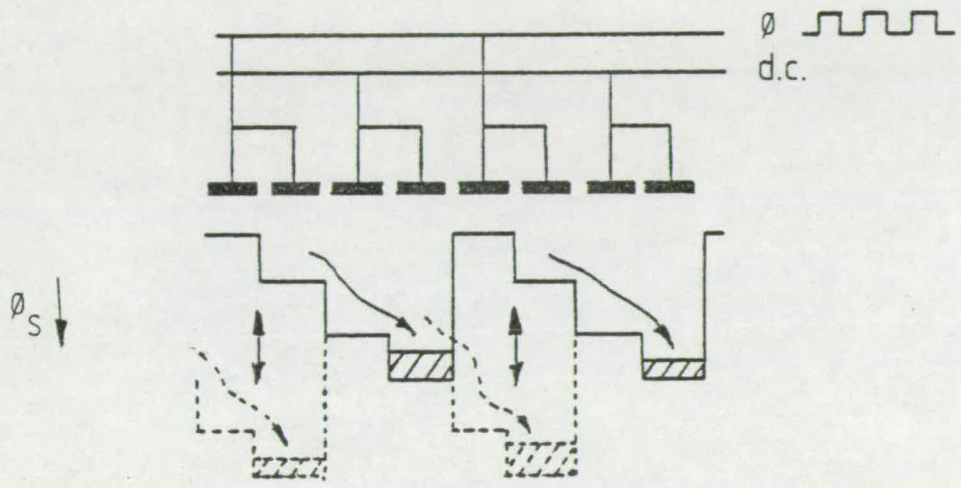


Figure 3.5: Single phase CCD

channel CCDs. Surface channel CCDs use a uniformly doped substrate so that charge packets reside at, and are transferred along, the semiconductor surface, where their potential minimum exists. Unfortunately, unsatisfied crystal lattice bonds exist at this surface, which can trap minority carriers and interfere with the charge transfer process.[22-24]

Now the potential minimum can be moved a short distance into the bulk of the substrate, so that charge packets avoid the surface entirely, by introducing a surface layer of opposite polarity to the rest of the substrate, using impurity diffusion or implantation.[36] The resulting structure is a buried channel CCD. Because the charge packets reside further away from the electrodes the charge storage capacity is reduced, and non-linearities are introduced into the charge-potential relationship. However, the transfer fringing field is enhanced, enabling faster operation.

For these applications the considerations of linearity, and particularly storage capacity are most important; surface channel devices are therefore used exclusively.

3.4 CHARGE INJECTION

Clearly some method of generating charge packets at the input of the CCD is required. Now for p-type substrates, minority carriers (electrons) can be generated at the semiconductor surface from a heavily doped n^+ region (called the input diode), to which electrical contact can be made. All electrical input techniques use such a diode to source charge into a metering well below an input electrode, such that the amount of charge present beneath the electrode is proportional to an input signal. This charge packet is then isolated and transferred down the CCD.

Several such input circuits have been devised,[12-15], and the two in most common use, namely 'diode-cut-off' and 'fill-and-spill', are illustrated in Figures 3.6 and 3.7. In the diode-cut-off scheme, Figure 3.6, the metering well is allowed to fill from the input diode until the surface potential has equilibrated. The charge packet is then isolated by a clocked input gate. The size of the charge packet is determined by the difference between the diode-forced surface potential, and the depth of the metering well when empty. The input signal may thus be applied either to the diode, or to the metering well electrode.

Charge packet size is set by the difference between two empty-well surface potentials in the fill-and-spill technique of Figure 3.7. Once this potential difference has been established, the input diode is briefly pulsed low, to *fill* both wells with charge. As the diode returns to its high level, charge *spills* from the two wells, leaving a trapped charge packet in the metering well. The input signal may be applied to either input electrode, with the other clamped at a suitable reference potential.

Both input techniques offer various degrees of linearity, depending upon the particular connection used. However, the most important aspect of a CCD system is not the input charge conversion characteristic, but rather the total characteristic from input to output.

Thus it is important to choose the input and output circuits in combination so that an overall linear transfer function is achieved, even if the charge conversion characteristics are non-linear.

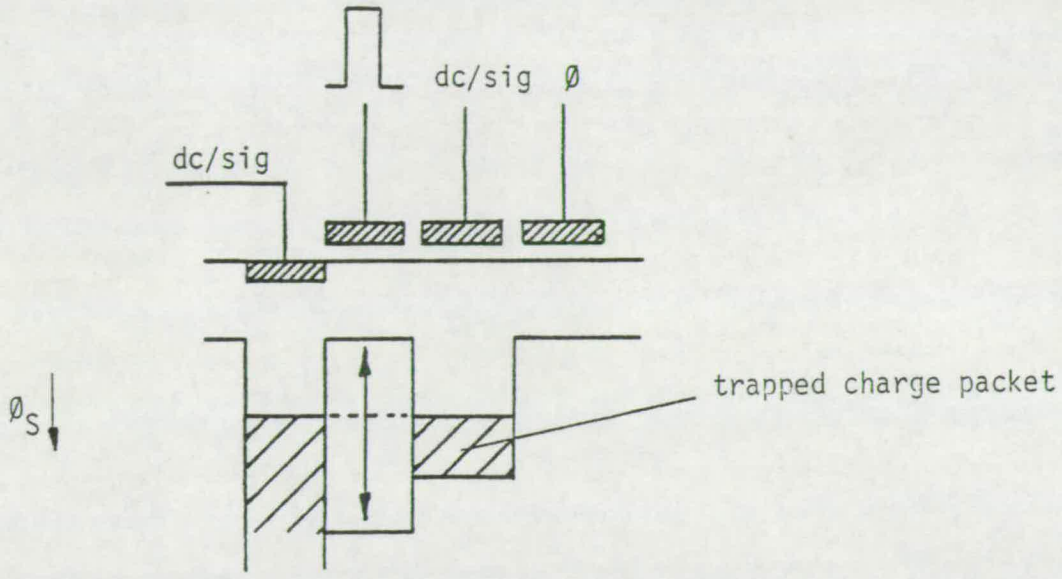


Figure 3.6: Diode-cut-off charge input technique

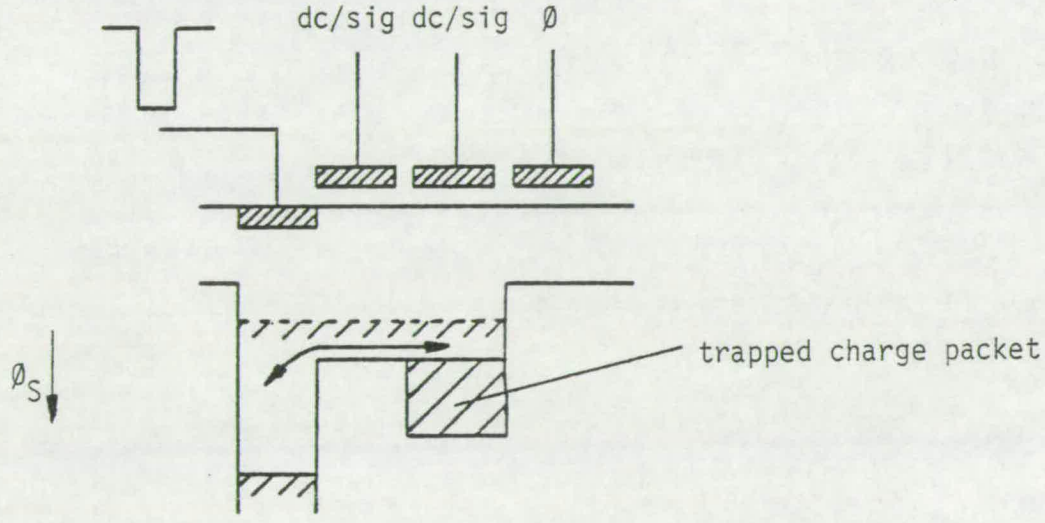


Figure 3.7: Fill-and-spill charge input technique

3.5 CHARGE DETECTION

Charge packets are normally sensed either through physical surface contact, via a heavily doped diode region, or else within the CCD structure itself, through the oxide capacitance beneath a special sensing electrode. The latter technique has a particular attraction for multiple tapped CCD delays, in that the structure of the CCD remains intact, so charge packets may pass along the device, yet be periodically sensed without interference to the charge transfer process.

In fairness this could also be true of the diode technique, with charge transfer into, and then out of, the diode in bucket-brigade[34] fashion. Howeverbucket-brigade charge transfer is less efficient and noisier than that within an intact CCD.

In Chapter 4 the theory of the floating-gate CCD tap is fully developed, with special regard to the realisation of high density tapped CCD delay lines.

At the end of a CCD delay line, charge packets are normally permanently removed through a single output diode clamped at a high potential, representing a charge sink.

3.6 CHARGE TRANSFER INEFFICIENCY

In a real CCD the charge packets are never perfectly transferred between locations; a small residual fraction is left behind. This smearing leads to a first order frequency-dependent attenuation, and account must therefore be taken of this effect in any CCD application.

Charge transfer inefficiency stems from two sources;

(i) Imperfect Free Charge Transfer. There are three basic charge transfer mechanisms;[9] drift due to a self-induced field, thermal diffusion, and drift due to the fringing field imposed at the semiconductor surface by the difference in potential between adjacent electrodes. These effects have been studied by several authors [18-21], and in general require sophisticated computer techniques[8,9] to form quantitative solutions, which lack generality. Clearly the removal of the charge is assymptotic toward the true empty well condition, so that truncating the transfer process at a finite time T results in a charge residual which can be expressed as a fraction, ε_f, of the original charge packet. This process sets an upper limit on clock frequency for a given maximum allowable value of ε_f. For surface channel devices with electrode lengths on the order of 10 μm, oxide thicknesses of 0.1 μm, and clock amplitudes of 10V, ε_f is typically 10^{-3} at $f_c \simeq 5$MHz.

(ii) Surface State Trapping. In surface channel CCD's signal electrons are stored at the semiconductor-insulator interface. This interface contains surface state traps, some of which are lower in energy than the free electrons, which therefore fill the exposed states and become trapped there. Now the filling time constant is known to be much faster than the equivalent emission constant.[22] Thus when a charge packet of any significant size occupies an area beneath an electrode, the surface states over that area are filled. When the charge packet leaves the electrode most of the trapped carriers remain, effectively to join

the following charge packet. A charge transfer inefficiency is induced because of the differing surface areas occupied by charge packets in a real potential well with sloping edges, as shown in Figure 3.8. Clearly the quantity of charge left behind at each transfer is directly proportional to the surface area occupied by the charge packet.

If the edges of the well are assumed to slope *linearly*, and if the CCD is operated with a fixed minimum charge level (fat zero), which always covers the surface states over an area A_0, then it can be shown[11] that the fractional residual at each transfer, the charge transfer inefficiency due to surface states, is a constant, ε_{ss}. As may be expected this constant is proportional to the significance of the variation in surface area, that is ε_{ss} is inversely proportional to the electrode length, and to the potential gradient at the edges of the potential well. For analogue CCD only the potential well edges perpendicular to the direction of charge flow, those between the electrodes, are significant because the aspect ratio of the electrodes is large.

Approximate expressions for ε_{ss} in terms of processing parameters may be derived, but when using a dedicated process only the general relationships discussed above are of concern. Such processes offer ε_{ss} typically in the region of 0.5×10^{-3}.

Because both charge transfer inefficiency terms are small, their combined effect may be given by

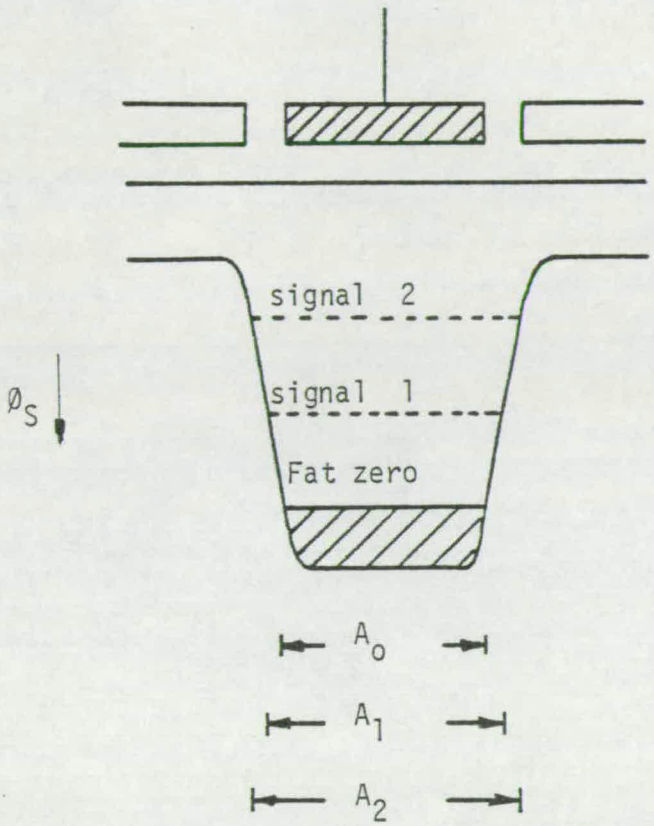


Figure 3.8: Variation in surface area occupied by differing charge packets in a potential well with sloping sides

$$\varepsilon = \varepsilon_f + \varepsilon_{ss}. \qquad (3.3)$$

Note that ε_{ss} is independent of clock frequency and thus imposes a lower limit on ε for low and medium clock frequencies, whereas ε_f will cause ε to deteriorate rapidly at high clock frequencies. Figure 3.9 demonstrates this variation of ε with clock frequency for a typical surface channel CCD. The corner frequency may lie between 1 and 5 MHz; above this the device rapidly becomes unusable. Buried channel CCD can push the corner frequency to 20 MHz or more, as well as reducing the minimum value of ε by avoiding the surface states altogether.

3.7 EFFECT OF CHARGE TRANSFER INEFFICIENCY ON FILTER PERFORMANCE

As an illustration of the effect of c.t.i, consider the time and frequency domain responses of a multiple stage CCD delay line. If the cumulative transfer inefficiency per stage, being p times that per transfer in a p-phase device, is ε and there are N stages, then Table 3.1 shows the distribution of charge within the delay line at different time frames after the introduction of a unit impulse at input stage 0, time 0. The table is built up by carrying forward a fraction $(1 - \varepsilon)$ of each charge packet in each time frame, and leaving behind the remaining fraction ε. The normalised impulse response observed (non-destructively) at stage n is given by the sequence in r defined by,

$$\frac{(n + r)!}{r!\, n!}\ \varepsilon^r (1 - \varepsilon)^n, \qquad (3.4)$$

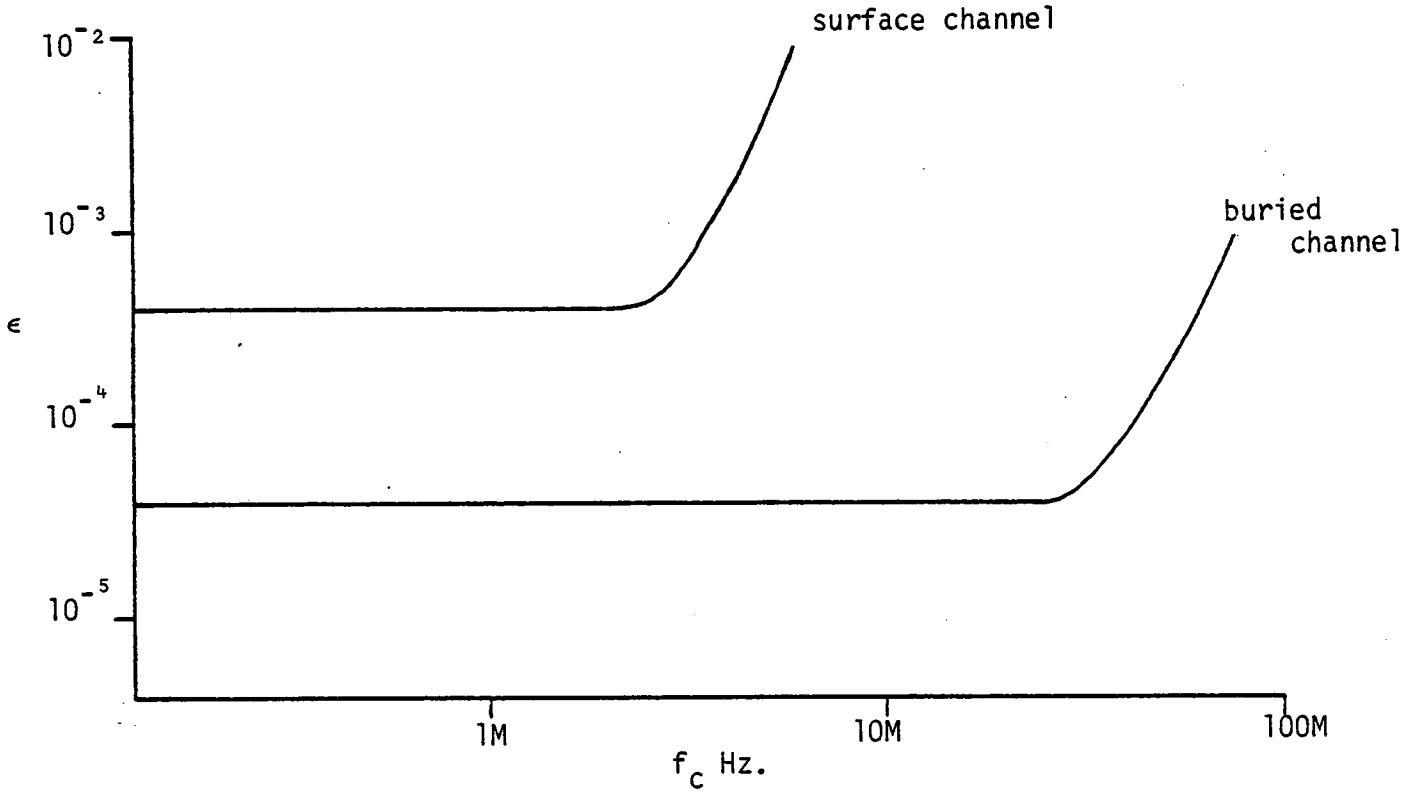


Figure 3.9: Variation of charge transfer inefficiency with clock frequency

Time Frame \ Stage	0	1	2	3	4
0	1				
1	ε	$1 - \varepsilon$			
2	ε^2	$2\varepsilon(1 - \varepsilon)$	$(1 - \varepsilon)^2$		
3	ε^3	$3\varepsilon^3(1 - \varepsilon)$	$3\varepsilon(1 - \varepsilon)^2$	$(1 - \varepsilon)^3$	
4	ε^4	$4\varepsilon^3(1 - \varepsilon)$	$6\varepsilon^2(1 - \varepsilon)^2$	$4\varepsilon(1 - \varepsilon)^3$	$(1 - \varepsilon)^4$

Table 3.1: Charge Distribution in a CCD Following a Unit Impulse Input at Stage 0, Time 0.

where r represents the residual number, r = 0 corresponding to the arrival of the first charge packet, r = 1 corresponding to the first residual, etc.

For small ε and large n (3.4) reduces approximately to

$$(1 - n\varepsilon)(n\varepsilon)^r, \tag{3.5}$$

giving an impulse response profile defined by the cumulative c.t.i. product $n\varepsilon$. For most practical applications $n\varepsilon < 1$ and the impulse response profile takes the form of an initial pulse, followed by a decaying tail of residuals. The c.t.i. product of a given CCD may in fact be estimated by measuring the ratio of the first residual to the primary response, giving $(n + 1)\varepsilon$ from (3.6).

The frequency response of the CCD is obtained by considering the n^{th} product of the z-transform of the impulse response of a single stage. Thus,[7]

$$F_N(z) = \left[\frac{1 - \varepsilon}{1 - z^{-1}}\right]^n$$

$$= z^{-n} \exp n(\ln[1 - \varepsilon] - \ln[1 - \varepsilon z^{-1}]) \tag{3.6}$$

The factor z^{-n} represents the expected delay and can be disregarded in such a discussion of signal degradation. The frequency response stems from (3.6) with the substitution[90]

$$z^{-1} = \exp(-2\pi i f/f_c), \tag{3.7}$$

which yields

$$|F_N(f/f_c)| = \exp\,[-n\varepsilon(1 - \cos 2\pi f/f_c)]. \qquad (3.8)$$

This expression gives the frequency response of the CCD a characteristic bell shape, the definition of which depends upon the product $n\varepsilon$.

These expressions for the time and frequency domain response of an inefficient CCD are verified in Figure 3.10, which shows impulse and frequency response measurements taken at two points on an experimental tapped CCD delay line (the design of which is reported later). The two impulse response photographs demonstrate the expected profile, and indicate cumulative c.t.i. products $n\varepsilon$, of 0.19 and 0.38 respectively. The theoretical frequency responses generated from (3.10) using these values agree well with the measured responses.

Signal degradation due to charge transfer inefficiency effectively limits the maximum filter length that may be used in any given application. In matched filtering applications, charge transfer inefficiency can cause a reduction in correlation peak sharpness and magnitude, as the 'match' becomes more imperfect. Analysis of this effect is complex, though computer simulation using (3.4) is feasible.

This limitation is considered in more detail in Chapter 9.

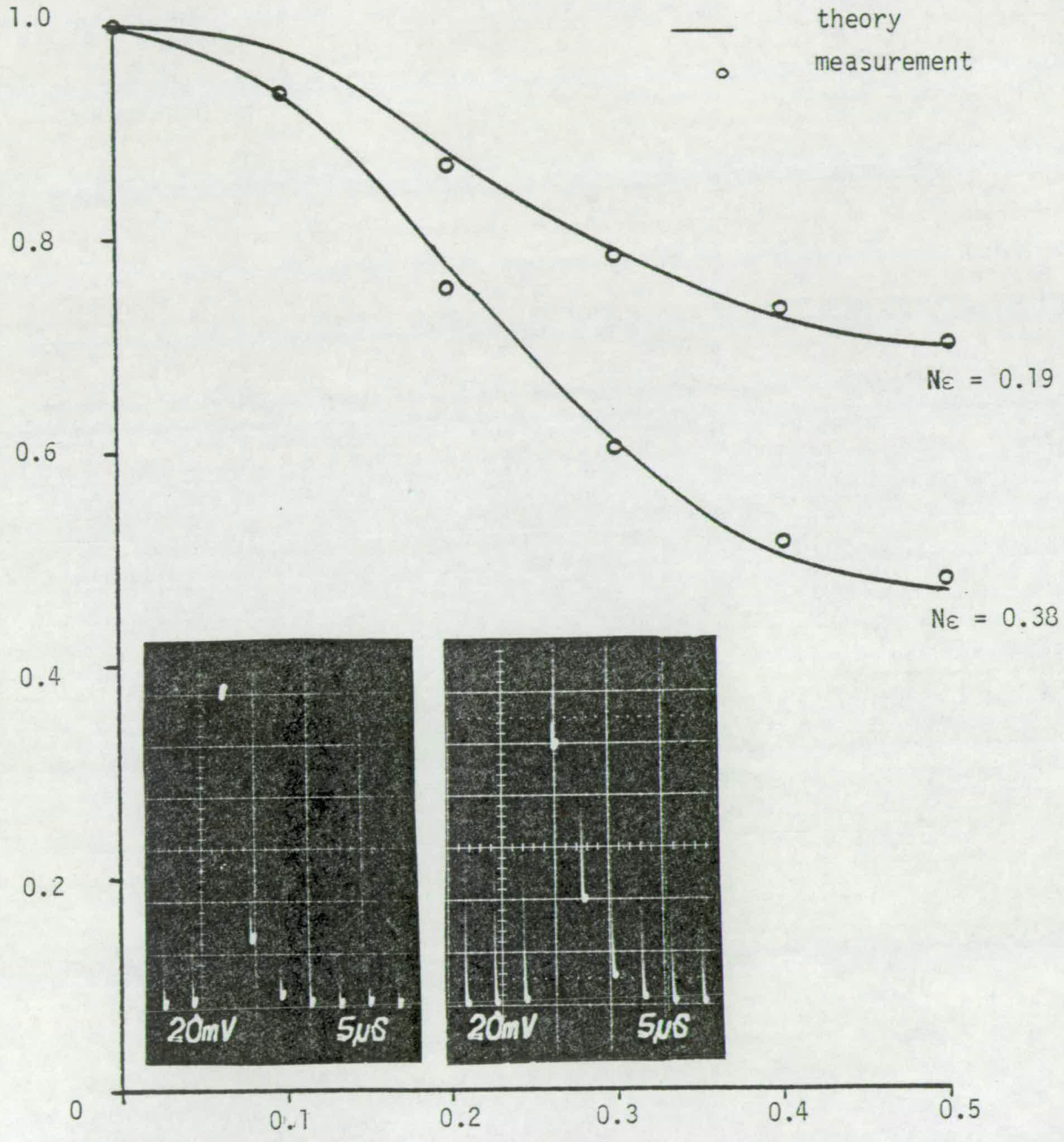


Figure 3.10: Impulse and frequency responses of an inefficient tapped delay line

3.8 MULTIPLE SAMPLING TECHNIQUES

For application of the CCD as a tapped delay line, a natural choice is to tap at every stage, as shown schematically in Figure 3.11(a). Now in an integrated filter realisation the cell circuitry might impose a minimum physical tapping pitch. This pitch may enforce CCD electrode lengths above those suitable for complete charge transfer at the maximum clock rate envisaged. Thus for operation at megahertz clock rates, surface channel electrode lengths should not generally exceed approximately 10μm. If the filter cell pitch implies electrode dimensions much longer than this, then it may be necessary to include more than one CCD stage between taps (and operate at a higher clock rate to achieve the same delay), as shown in Figure 3.11(b).

Including more transfers between each tap would appear to increase the c.t.i. induced at medium clock frequencies by surface trapping, both because of the increased significance of the edge effect with shorter electrodes, and because of the increased number of transfers. However, if the signal bandwidth is still determined by the tap delay, some oversampling redundancy is present and this serves to reduce the effect.

There are several ways of exploiting oversampling redundancy[91,92], but the most efficient is the multiple sampling technique illustrated in Figure 3.11(b) for two-stage redundancy. One input signal sample is held and injected into the delay line in two identical successive

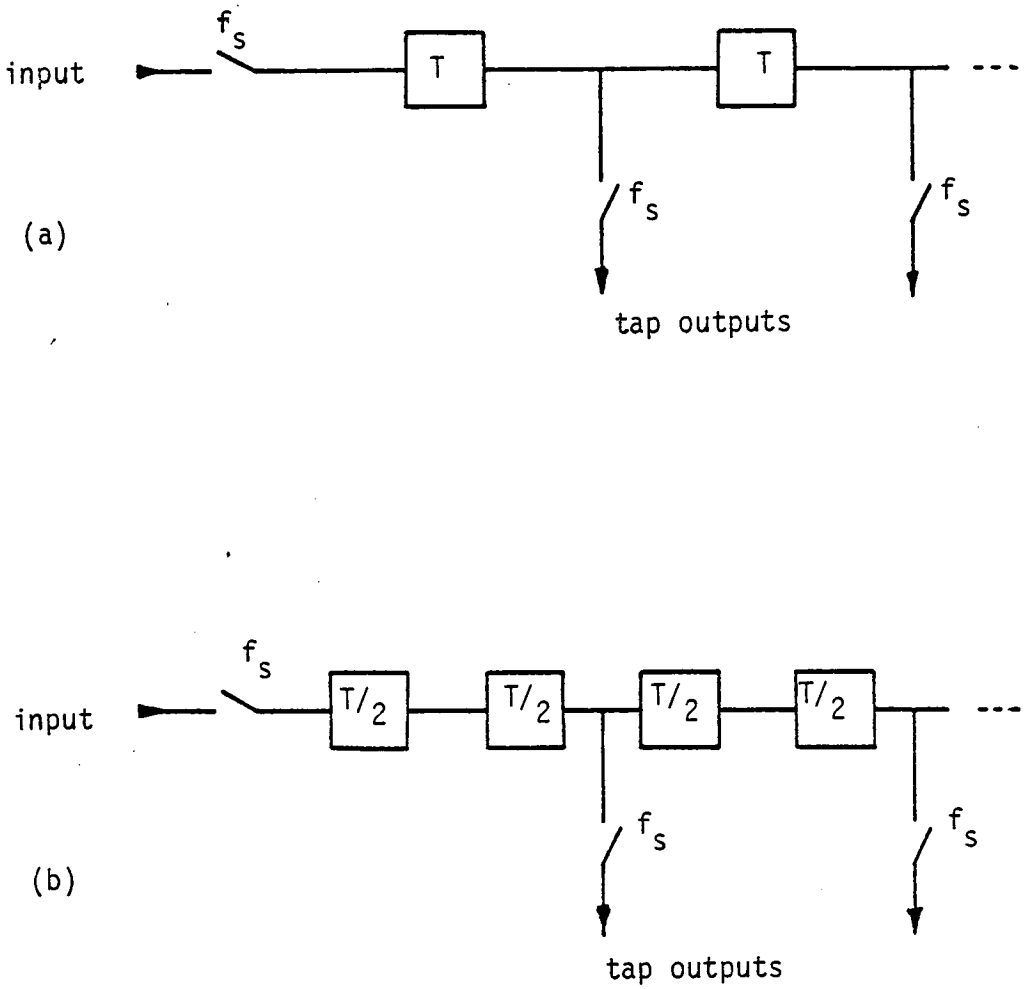


Figure 3.11: Tapped CCD sampling schemes

(a) conventional operation

(b) double-sample operation

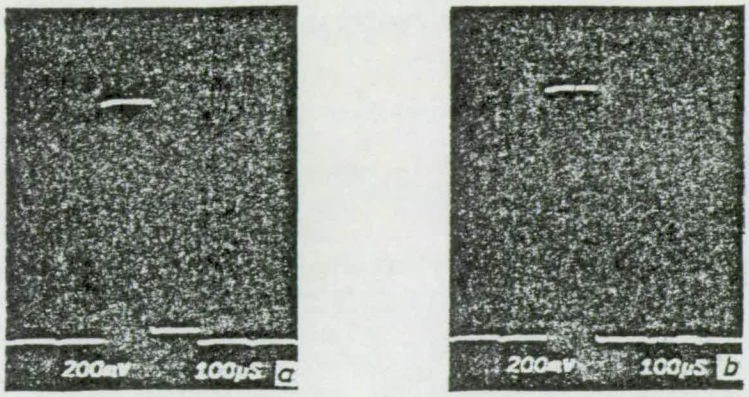
charge packets, and the second charge packet of each pair is sampled at the tap outputs. The charge loss of the second packet by transfer inefficiency is compensated by the (ideally) identical loss of the leading packet during each transfer. The leading charge packet also blocks the residual effect of preceding signal samples.

The impulse response of the double-sampling scheme is obtained by considering alternate terms of the conventional response (3.4) to two adjacent impulses, and allowing for 2n stages between n taps[92];

$$\frac{(2n + 2r)!}{2r!2n!} \left(1 + \frac{2n\varepsilon'}{2r + 1}\right) \varepsilon'^{2r} (1 - \varepsilon')^{2n} , \qquad (3.10)$$

where ε' is the new c.t.i. per reduced-length CCD stage.

Figures 3.12(a) and (b) illustrate the beneficial use of the double-sampling scheme on a tapped delay line. The first trace shows the impulse response of a conventionally operated CCD after 26 three-phase stages, clocking at 10 kHz. The second trace shows an improved impulse response from the same device after 52 stages, using a double-sampling scheme at 20 KHz. The total delay is the same in both cases, but the response of the double-sampling scheme is clearly superior.



(a) (b)

Figure 3.12: Transfer inefficiency improvement using double sampling

(a) conventional impulse response after 26 CCD stages @ f_c = 10 kHz

(b) impulse response after 52 CCD stages using double sampling technique @ f_c = 20 kHz, post tap sampling @ f_c = 10 kHz.

Other techniques which exploit redundancy are possible, such as alternate-zero sampling and continuous double-rate sampling[92], but neither of these betters the response of (3.10). Thus where multiple CCD stages must be included between taps to ensure high frequency operation, a multiple sampling technique should be used to obtain the best impulse response.

3.9 DARK CURRENT

In addition to charge transfer inefficiency, the CCD suffers one other primary limitation, involving leakage of thermally generated minority carriers into the signal charge packets. Thermal excitation generates electron-hole pairs in the depletion regions beneath the CCD electrodes. The minority carriers are attracted to the potential minimum at the semiconductor surface, where they are indistinguishable from the signal charge. In a continuously clocked CCD this thermally generated charge results in a quiescent 'offset', which builds up uniformly along the delay line and limits the time a signal may stay in the device before significant degradation occurs.

The amount of dark current charge collected by stage n is,

$$Q_{dark} = \frac{nJ_D A}{f_c}, \qquad (3.11)$$

where A is the total area of one CCD cell, and f_c is the clock frequency. J_D is the dark current surface density which is exponentially temperature dependent (through the intrinsic equilibrium carrier density), and typically has a value between 5 and 50 nA/cm^2, limiting total storage

times to the order of 1 second.

Therefore, while c.t.i. limits the total number of CCD stages that may be used, dark current limits the total time that may be spent in any one device. Both processes are noisy and thus further limit device performance. This aspect is treated in Chapter 9.

Note that the quiescent dark current offset can be blocked between cascaded devices by a.c. coupling. Thus it is only the total time within any one device that is limited.

3.10 THE MOS TRANSISTOR

The upper electrode of an MOS capacitor can be used to invert and control a conduction channel between two heavily doped (source and drain) regions as shown in Figure 3.13(a). Such an arrangement is termed an MOS transistor, having source, drain and gate connections. Figure 3.13(b) shows typical simplified channel current characteristics as a function of the controlling gate voltage. Note that, for a given gate voltage, the channel current initially rises for increases in drain-source voltage. Above a certain point however there is no further increase in channel current and the device is said to be saturated. Below a certain gate-source voltage, known as the threshold voltage , no channel current flows at all since the channel is no longer inverted and the device is said to be cut-off.

The first order equations which describe the characteristics of Figure 3.12 are (after Crawford[6]);

In region 1: $$I_{DS} = \beta_o \frac{W}{L} \left([V_{GS} - V_{T_o}] V_{DS} - \frac{V_{DS}^2}{2} \right) \qquad (3.12)\ (a)$$

In region 2: $I_{DS} = \frac{\beta_o}{2} \frac{W}{L} (V_{GS} - V_{T_o})^2$ (3.12)(b)

where $\beta_o = \mu C_{ox}$ (a process dependent gain constant)

W = Transistor width

L = Transistor length (source to drain)

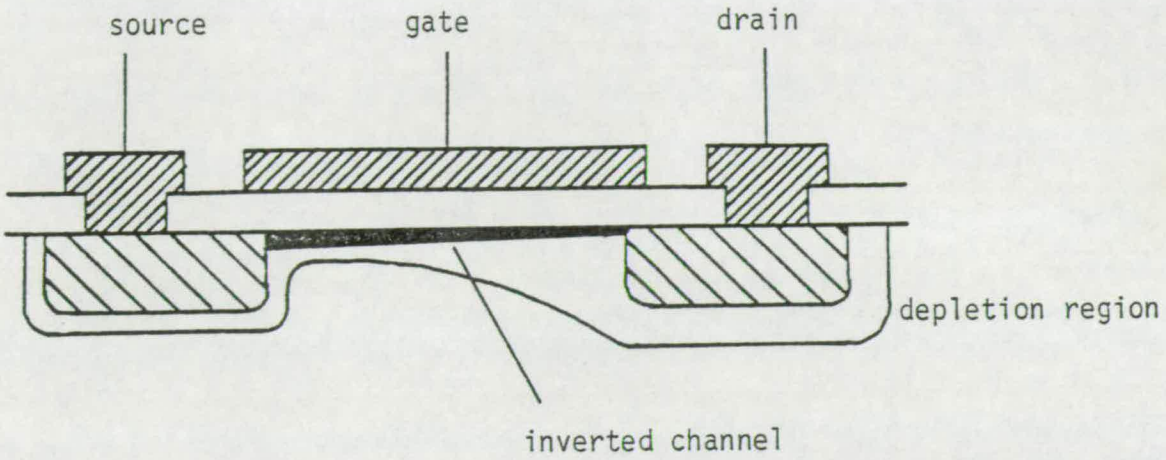
V_{T_o} = Threshold voltage.

Clearly the 'gain' of a transistor may be controlled by the width:length (or aspect) ratio when the geometry of the device is defined at the layout stage. Typical values of β_o for oxide thicknesses of 800 Å, are 25 $\mu A/V^2$, and V_{T_o} may be varied according to requirements by locally controlling the substrate doping density N_A. Devices having several different values of V_{T_o} may be achieve on one chip by using different masks to define different (ion-implanted) doping levels. Processes may offer one or more standard values to be used. Typical values range from -10V to +10V.

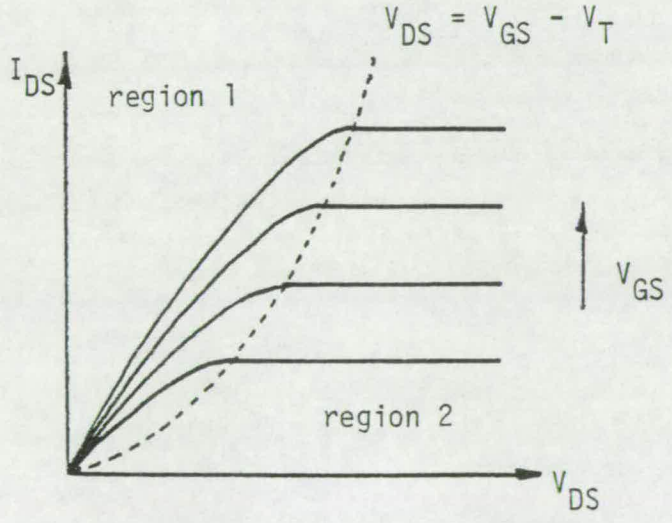
When designing linear MOST circuitry the simplified characteristic equations are often insufficient, and it becomes necessary to consider three dominant second-order effects.

(a) Body effect. The threshold voltage is not constant, but becomes more positive as the source-substrate bias V_{SB} is increased. Thus the true threshold voltage is given by[4];

$$V_T = V_{T_o} + \gamma \sqrt{V_{SB} + 2\phi_F} \qquad (3.13)$$



(a) Cross-section



(b) Simplified characteristics

Figure 3.13: The MOS transistor

where $\gamma = \sqrt{2N_A q \varepsilon_o \varepsilon_s}/C_{ox}^2$ which is typically in the range 0.3 - 1.3, and $2\emptyset_F \simeq 0.6V$ in silicon.

(b) Channel length modulation. The drain current does not remain constant in saturation, but increases slightly with drain voltage. This occurs because the width of the depletion region around the drain increases with the drain voltage, so shortening the channel length L. This effect modifies the drain current expressions by a factor[2]

$$\frac{L}{L - \sqrt{2\varepsilon_s(V_{DS} - V_{DS_{sat}})/qN_A}}, \qquad (3.14)$$

where L = real channel length. This induces a finite small signal drain conductance in saturation, which is minimised for long channels. Thus stages which rely on a saturated MOST current source may deliberately use a long channel geometry.

(c) Field dependent mobility. The mobility, μ, of minority carriers in the MOST channel is not constant, but depends inversely upon the gate-induced electric field. To first order this modifies the gain constant β_o to[6];

$$\beta = \frac{\mu_o C_o}{1 + \theta(V_{GS} - V_T)} \qquad (3.15)$$

where θ is an empirical constant, typically in the range ± 0.05.

3.11 LINEAR MOST CIRCUITS

As well as its common switching application in digital circuit realisations, the MOST is a useful linear component,[32] finding application within buffer stages and inverting amplifiers (which may ultimately form operational amplifiers), and also as a linear multiplying element. Two of these circuits, the high input impedance source follower buffer, and the high gain inverter are analysed in Appendix II . Both find common use within the filter realisations described here. The more sophisticated realisations, of multipliers and operational amplifiers, are the individual subjects of subsequent chapters.

CHAPTER FOUR : CCD TAPPED DELAY LINES

The CCD offers great potential in forming a compact analogue delay element. However if it is to be useful in realising the programmable filter, then a method of achieving multiple staggered-delay tap outputs is required. Furthermore, if a high integration density is to be achieved, this method must imply a minimum of silicon area per output,and be easily expandable to any required number of taps.

Equally crucial to the success of such a scheme is the availability of a complementary input technique which will result in a linear overall transfer function, from signal input to tap output. The theory and practice of such a tapping scheme and an associated input technique are fully developed in this Chapter.

A suitable tapping technique, involving non-destructive sensing of charge along a single CCD channel, is that of Floating Gate Reset (FGR) tapping. This scheme, originally suggested by MacLennan[43,11], allows a single charge packet passing along a CCD to be repeatedly sensed without obstructing its passage. FGR tapping is elegant and compact, the layout area increasing linearly as more points are added, by simply extending the single CCD channel.

Examination of charge injection techniques shows that the diode-cut-off scheme[12] is a most suitable complement to the FGR tap, providing a simple, edge-sampling signal input. When combined with a dummy tap, this input technique offers the desired ideal linear transfer function from signal input to FGR tap output.

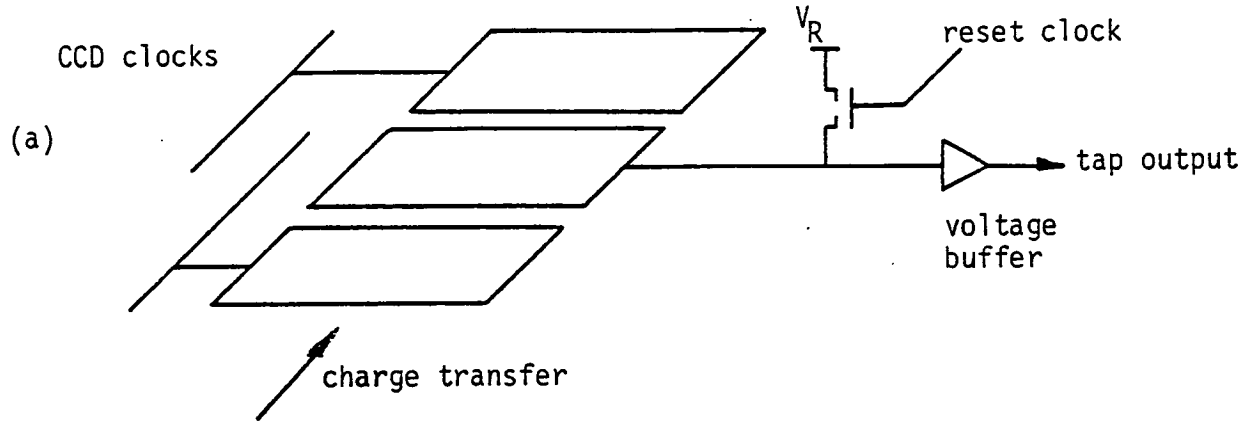
4.1 FGR TAPPING FUNDAMENTALS

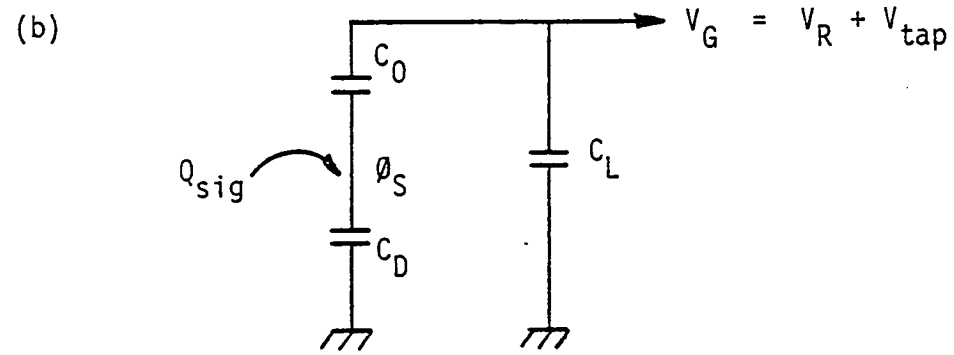
A simplified diagram of a floating gate reset (FGR) tap is shown in Figure 4.1. The tap itself is a normally formed electrode within the CCD channel, but it is not connected to any of the clock phases. Instead, whilst the potential well under the electrode is empty of charge, the tap is reset to a reference voltage, V_R, and is then isolated, or allowed to 'float'. Any charge which is subsequently injected underneath this tap electrode is sensed through the oxide capacitance, stimulating a charge redistribution within the tap structure and causing a change in the tap potential, in some proportion to the magnitude of the charge packet injected. This tap voltage is buffered to provide a signal output at low impedance, using a MOST source-follower.

Although the charge packet transduces a voltage change in the tap, it still remains within the CCD channel and is available for further transfer. Clearly when the charge packet is removed, a reverse process takes place, and the tap must return to its original reference potential if the principles of charge conservation apply. However, gradual charge leakage from the sensing node is likely to occur, and so the tap must periodically be reset.

In practice the CCD clocks are used to inject, and then to remove the charge packets to and from the tap, in a ½ phase operation. It is convenient also to use one of the available CCD clocks for the reset operation. Details of these modes of reset, injection, and removal are considered in the following sections.

A capacitive model of the tap structure for a surface-channel CCD is shown in Figure 4.1(b), where the total depletion and oxide capacitances underneath the tap electrode are given by C_D and C_O





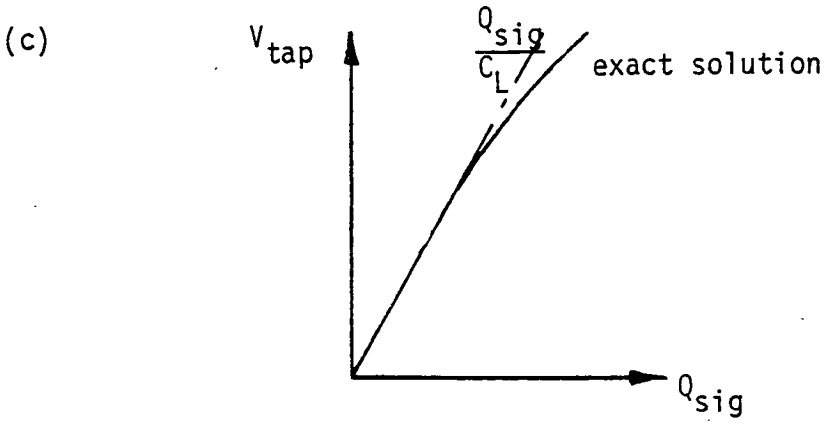


Figure 4.1: FGR tap principles

(a) Schematic
(b) Capacitive model
(c) Variation of tap potential with injected charge

respectively, and C_L represents the total load on the tap external to the CCD channel, including strays and buffer input capacitance. A rigorous solution for the change in tap potential, V_{tap}, upon the introduction of a charge packet, Q_S, is complicated by the variation of the depletion capacitance with surface potential, ϕ_S. However, such a solution has been obtained by MacLennan[11], who considered the characteristic equations;

$$V_{tap} = \frac{Q_S C_O}{C_O C_D + C_D C_L + C_O C_L} \tag{4.1}$$

$$C_D = C_O \checkmark \frac{V_A}{2\phi_S} \tag{4.2}$$

$$\phi_S = V_G - V_{FB} + \frac{Q_S}{C_O} + V_A - \{2V_A[(V_G - V_{FB}) + \frac{Q_S}{C_O}] + V_A^2\}^{\frac{1}{2}} \tag{4.3}$$

Their simultaneous solution gives a 6^{th} -order polynomial in V_{tap} that may be solved numerically, to give a transfer function of the form shown in Figure 4.1(c). Also shown is the limiting sensitivity relationship,

$$V_{tap} = \frac{Q_S}{C_L} \quad , \tag{4.4}$$

which is a good approximation to (4.1) for $C_D << (C_O \text{ or } C_L)$. This inequality is improved when the substrate doping is small, or when a negative substrate bias is applied.

The inclusion of a non-clocked electrode within the CCD structure can reduce its charge-handling capability, and the connection of that electrode in floating gate mode will exaggerate this effect. Care must therefore be taken that a useful dynamic range is obtained, in

terms both of the magnitude of the tap output voltage, and of the size of the charge packet itself.

Consider the surface potential underneath the tap as a function of the tap voltage, as shown in Figure 4.2. In general the useful surface potential range will be limited, by conditions applying to the CCD electrodes on either side of the tap, to a maximum value ϕ_{max}. Now assume that a charge packet $\hat{Q}_{tap}$ fills this well, then from the capacitive model of Figure 4.1(b),

$$\hat{Q}_{tap} = \phi_{max} \left(C_D + \frac{C_O C_L}{C_O + C_L}\right) \tag{4.5}$$

$$\simeq \phi_{max} \frac{C_O C_L}{C_O + C_L} . \tag{4.6}$$

Note that the capacitive multiplying factor in (4.6) is simply C_O for non-floating electrodes. The reduced charge-handling capability of the floating electrode stems from a feedback effect between the electrode potential and the surface potential below it. As electrons are injected underneath the electrode, its potential is lowered and thus the effective potential-well depth is reduced.

Combining (4.4) and (4.6) gives,

$$\hat{V}_{tap} \simeq \phi_{max} \frac{C_O}{C_O + C_L} . \tag{4.7}$$

Clearly minimising C_L will maximise the tap output voltage, but this may reduce $\hat{Q}_{tap}$ unacceptably through (4.6). Conventionally dynamic range may always be improved by increasing the width of the CCD channel, and hence its charge-carrying capacity in proportion; the dominant charge noise term (discussed in Chapter 9) increases only as the square

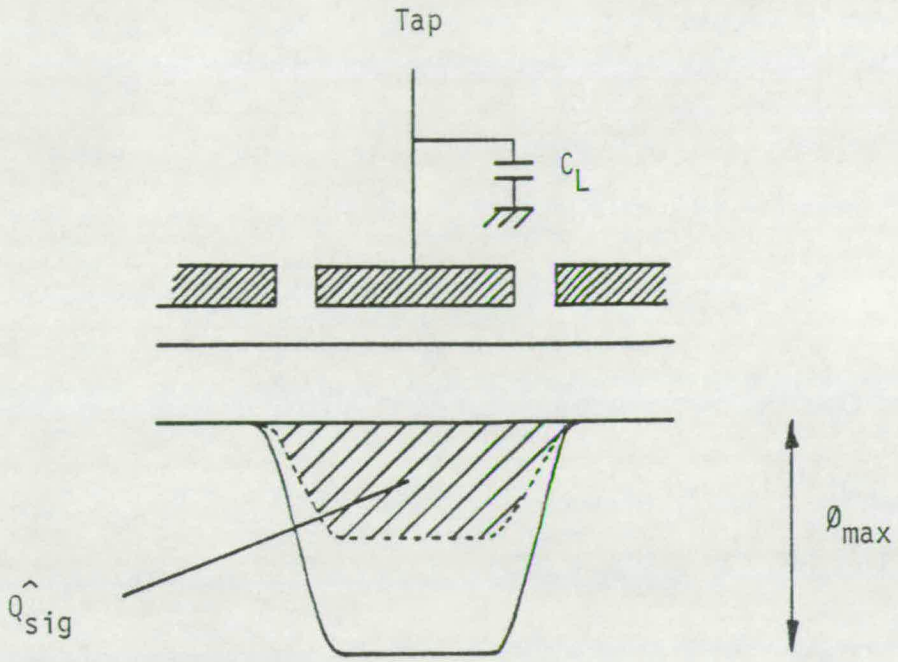


Figure 4.2: Maximum charge handling condition for a floating-gate tap

root of gate area. However, if $C_0 >> C_L$ for a tapped delay line, then the maximum size of the charge packet may be fixed through C_L (4.6), and it can become detrimental to further increase C_0.

Given sufficient noise data on the process concerned some optimum theoretical compromise between C_0 and C_L may be found. Such information is often not available however, and in any case C_L and C_0 are practically subject to minimum and maximum limits from consideration of the topology and area of the layout. To simplify the design process for application of such a multi-tapped delay line in a high-density PTF, a useful practical condition is:

$$C_L \simeq C_0. \tag{4.7}$$

One further point must be made concerning the use of FGR taps within CCD delay lines, relating to their effect on charge-transfer inefficiency. At low clock frequencies, well below the 'corner' at which free charge transfer effects dominate, fast-interface-state trapping limits c.t.i. to approximately 0.5×10^{-3} per transfer for surface channel devices.[35] These effects are not well understood but we may conclude that, where the inclusion of the tap results in a weakened fringing field between electrode edges, the c.t.i. may be adversely affected. No practical comparison is available, but measurements on processed filters demonstrate average values of c.t.i. still within the region of 0.5×10^{-3} per transfer, allowing suitable operation in all of the applications envisaged.

4.2 DIODE - CUT - OFF INPUT

Although the intrinsic charge-voltage transfer characteristic of the FGR tap is not linear (Figure 4.1(c)), an overall voltage

transfer function which is so may be obtained by employing a suitable input technique. Consider initially a tap structure with a constant load capacitance C_L. It can be seen from the capacitive model of Figure 4.1(b) that a linear relationship does exist between a change in surface potential, and the corresponding signal appearing at the tap gate, such that

$$V_{tap} = \frac{\Delta\phi_S \, C_0}{C_0 + C_L} , \tag{4.9}$$

where $\Delta\phi_S$ is the applied change in surface potential. By using an input method which sets the surface potential directly (e.g. the diode-cut-off technique)[12], and an FGR tap structure in place of the metering well, as shown in Figure 4.3, it is thus possible to obtain a linear input/output relationship at the metering tap. The amount of charge trapped will not be a linear function of the signal but, when transferred along the register, it will give rise to identical linear output signals that are an exact replica of those at the metering tap. Care must be taken that the electrical conditions present at the metering tap are identical to those at subsequent output taps.

Operation of the diode-cut-off input stage is shown in Figures 4.3 and 4.4. Whilst the clock applied to the electrode following the tap is off, an input gate is opened and the surface potential underneath the metering electrode is set by charge flooding from the input diode. Once the surface potential has settled, the input gate is shut off and the resultant trapped charge packet, representing the required signal sample, is available for transfer. Note that the input signal is effectively edge-sampled, avoiding the requirement

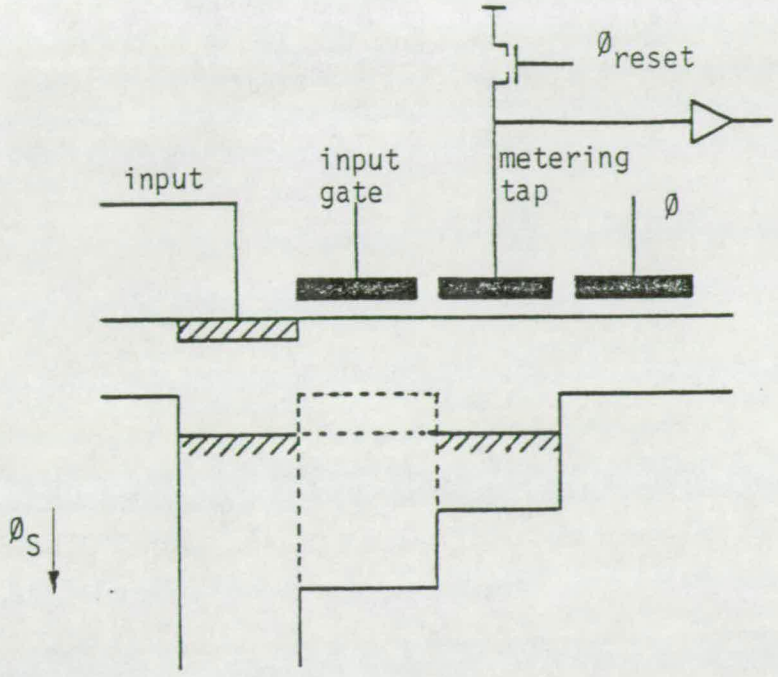


Figure 4.3: Diode-cut-off input structure

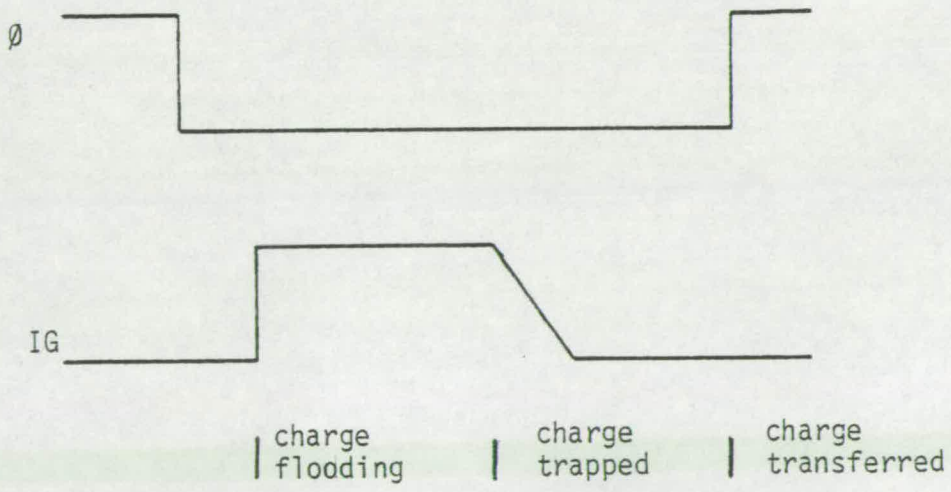


Figure 4.4: Diode-cut-off timing

for any pre-input sample-hold stage. The gain of the voltage transfer function is given by

$$A_V = \frac{C_O}{C_O + C_L}, \qquad (4.10)$$

which is always less than unity, and may be closer to 0.5.

In general the load capacitance C_L is not ideally constant, nor is the voltage buffer following the tap ideally linear, however for normal signal amplitudes and processing conditions neither effect distorts the output signal by more than approximately .1% (-60dB). Partitioning of the residual charge underneath the closing input gate is a possible problem however. According to the rate of this falling edge, a proportion of the charge will be pushed forward into the metering well, resulting in a non-linear error in the signal sample. This effect may be minimised by using a slow falling edge, or by reducing the size of the input gate relative to that of the metering well.

4.3 FGR TAPPED DELAY LINE EXAMPLE I

A photomicrograph of an experimental metal-gate FGR-tapped CCD delay line section is shown in Figure 4.5. This register is used in a 64-point PTF realisation which is described later. To obtain an optimum filter layout, taps from the CCD must emerge at a pitch of minimum dimension suitable for integration of the remainder of the filter cell circuitry. On the SET process used for this realisation the minimum filter cell pitch achievable is 56μm.

Such a metal-gate process offers a 3-electrode CCD cell for 3-phase, or $2\frac{1}{2}$ phase operation. In this case two CCD cells may be

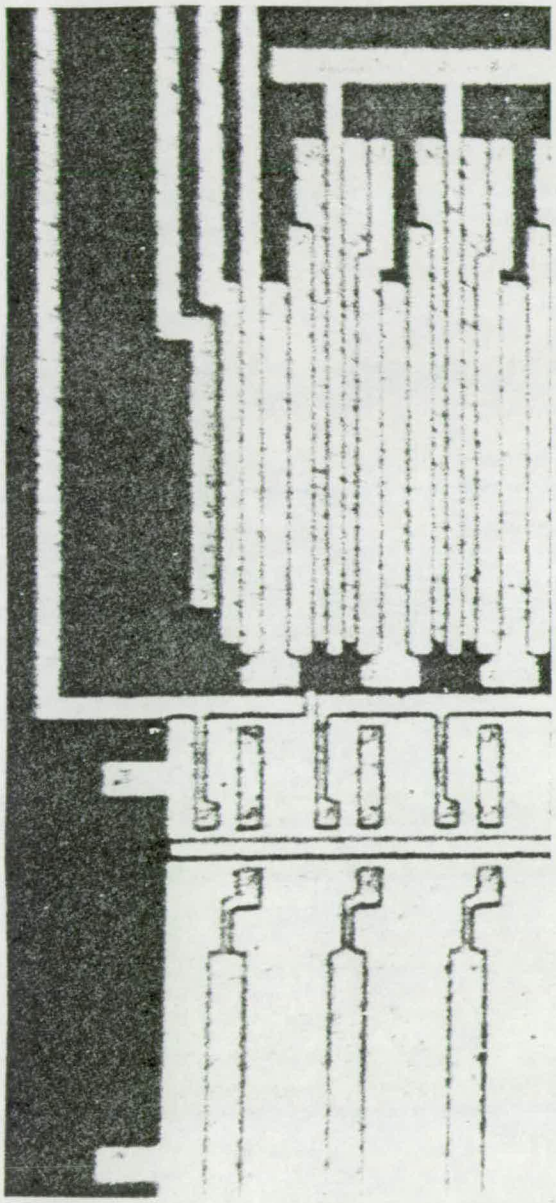


Figure 4.5: Photomicrograph of input stage and initial sections of a metal-gate FGR tapped CCD delay line

incorporated between taps to yield a suitable average electrode length of 9 μm. The alternately tapped structure permits a reduction in c.t.i. effects, either by over-sampling the input signal at twice the necessary frequency, or by using a double-sampling scheme and monitoring only alternate samples at the tap outputs.[92]

Each cell of the tapped delay line, shown schematically in Figure 4.6, comprises two 3-electrode CCD stages; the first being fully clocked 3-phase, and the second being 2½-phase clocked with the third electrode isolated as a floating-gate tap. The tap incorporates a reset transistor M1, and a source follower buffer M2, M3.

Electrode lengths represent final values in μm, allowing for 1 μm interelectrode gaps. These lengths were ratioed in an attempt to maximise charge transfer efficiency, by making longest those electrodes with the weakest fringing fields.[93] The CCD width was chosen to satisfy the $C_0 = C_L$ compromise. In this technology, where CCD electrodes are well isolated, the tap load capacitance C_L is dominated by the input capacitance of the tap buffer, and by stray capacitance outside the CCD channel. For this configuration then;

$$C_0 \simeq C_L \simeq 0.3 \text{ pF}. \tag{4.11}$$

These values give the following tap sensitivity relationships;

$$\frac{\partial V_G}{\partial \phi_S} = \frac{C_0}{C_0 + C_L} = 0.5 \tag{4.12}$$

$$\frac{\partial Q_S}{\partial V_G} \simeq C_L \simeq 0.3 \text{ pC/volt}. \tag{4.13}$$

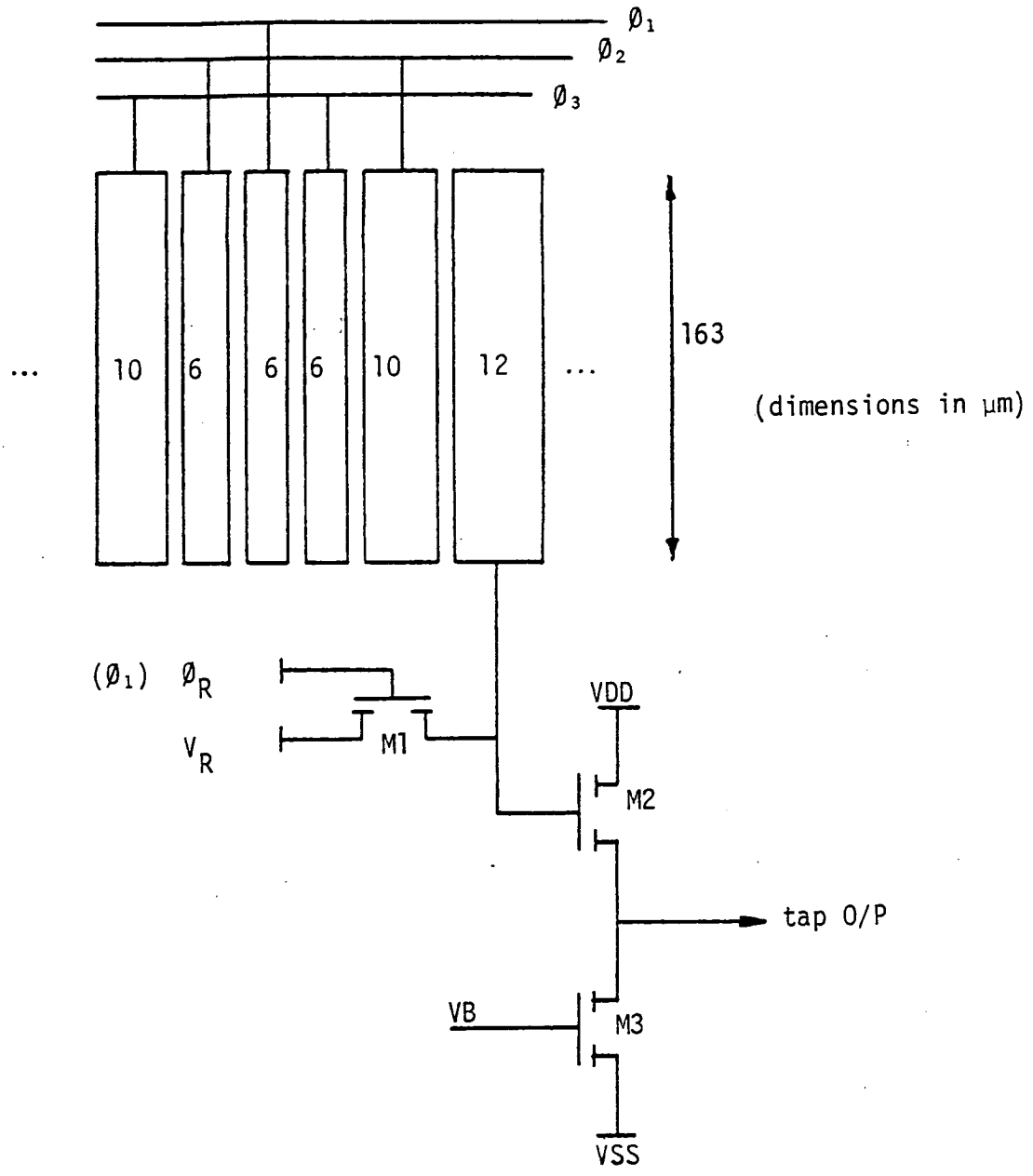


Figure 4.6: Cell schematic of a metal-gate FGR tapped CCD delay line

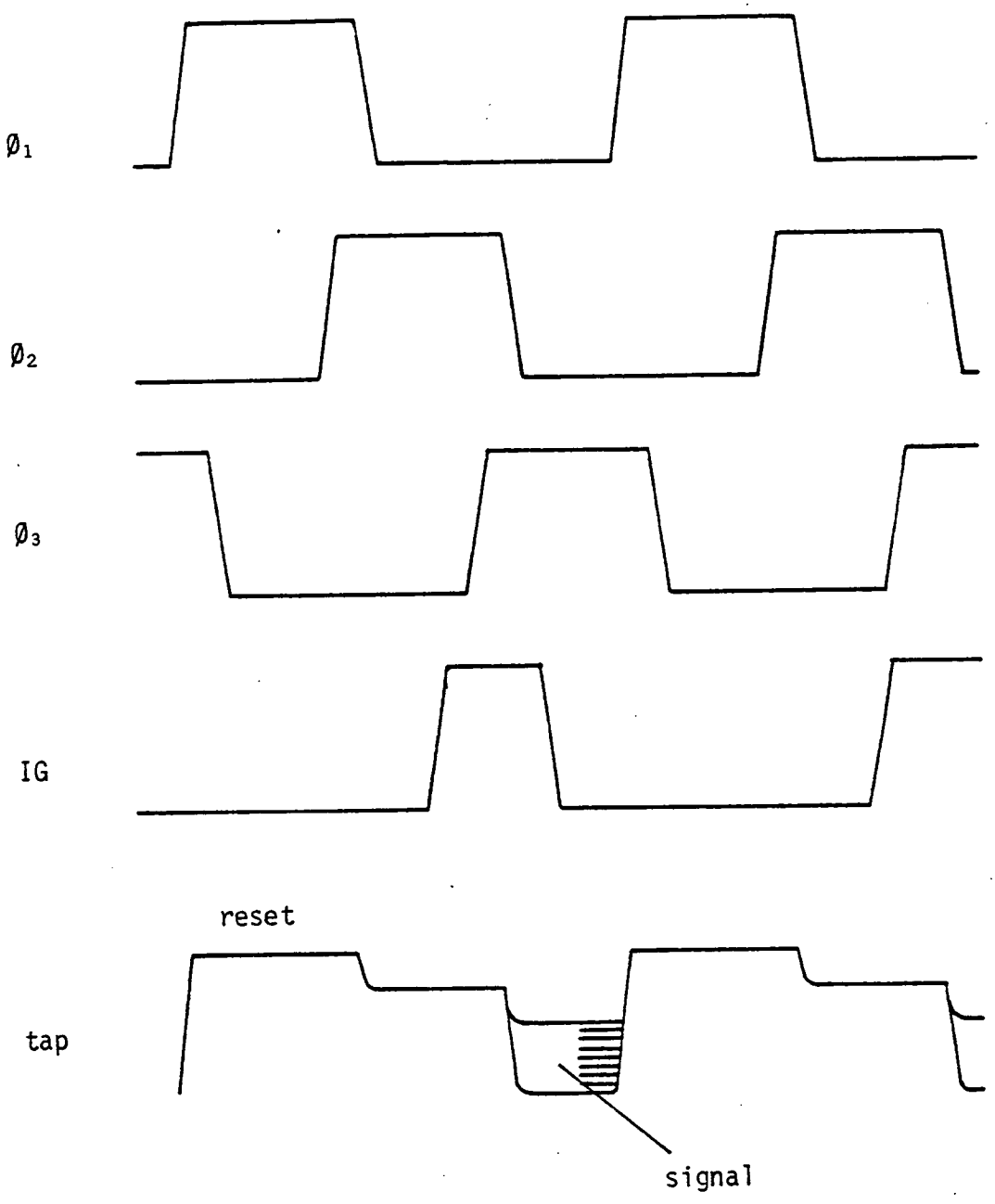


Figure 4.7: CCD and tap output timing diagram.

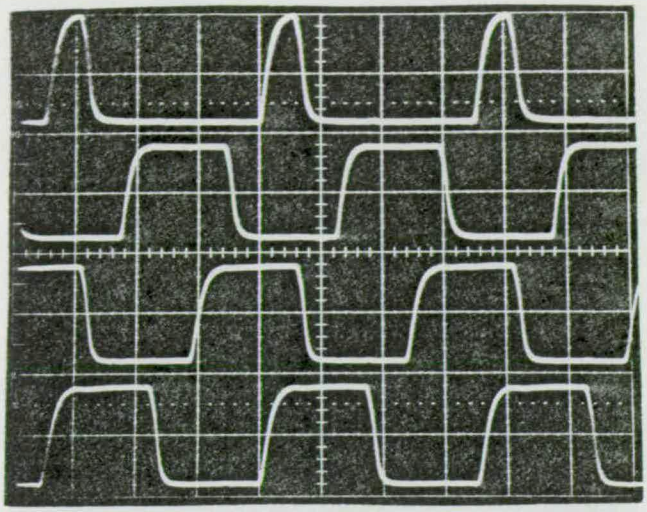
Figure 4.7 shows a timing diagram for the CCD and the tap output waveform. Clock phase $\emptyset_1$ is conveniently used to reset the tap, since it complies with the conditions set out in the previous section. Thus the expected tap output signal, as shown in the figure, comprises a reset period ($\emptyset_1$), a period of isolation ($\emptyset_2$), and then a period in which the charge transferred underneath the tap is detected. Because the tap becomes isolated there is some chance of capacitively-coupled breakthrough on the falling edges of clock phases $\emptyset_1$ and $\emptyset_2$. The electrodes are well isolated in this technology however, so the effect is slight.

A diode-cut-off input scheme, identical to that previously discussed and shown in Figure 4.3, is used. However, because none of the three overlapping phases qualify as an input gate pulse, an additional input phase is required, as included in Figure 4.7.

This FGR-tapped metal-gate CCD delay line operates successfully, both in discrete form as a 32-point TDL, and as a 64-point TDL in an integrated programmable filter. Figure 4.8 shows three oscilloscope photographs demonstrating performance. The photographs show; the complete set of clock waveforms used (c.f. Figure 4.7), the delaying of a sinusoidal waveform (3v p-p), and detail of the tap output waveform, also as shown in Figure 4.7.

The peak linear output signal attainable with these devices was limited by gross distortion effects, which are attributed to spurious surface potential barriers between the electrodes, as discussed by Hobson.[8] Because the gaps between the electrodes are not electrically controlled this has a variable effect and leads to non-uniform performance from device to device. Typical measurements for these devices give c.t.i. at 5.10^{-4} per transfer, and total harmonic

(a)



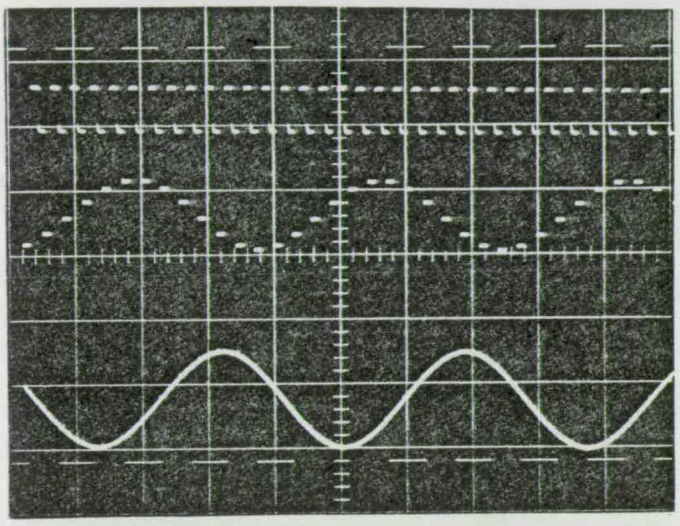
INPUT GATE

$\emptyset_1$

$\emptyset_2$

$\emptyset_3$

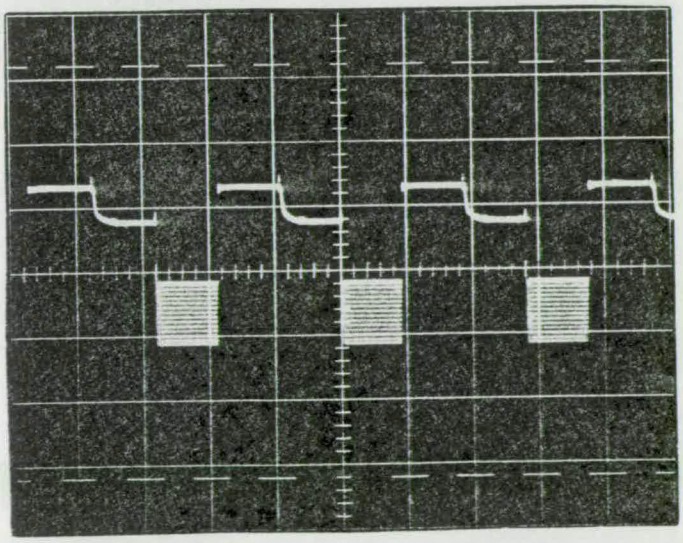
(b)



TAP OUTPUT

DIODE INPUT

(c)



TAP OUTPUT

Figure 4.8: Operation of metal-gate FGR tapped delay line

(a) Clock waveforms. X 200nS/div Y 20v/div
(b) Delay of Sinusoid. X 50µS/div Y 1v/div and 2vv/div
(c) Tap waveform detail. X 5µS/div Y 1v/div

distortion at -34 dB for output signals of 2V p-p, using clock amplitdues of 27V.

4.4 FGR TAPPED DELAY LINE EXAMPLE II

A second FGR-tapped CCD delay line, for use in a later PTF realisation, has been developed in a double-level polysilicon gate technology. A section of the input and first taps is shown in the photomicrograph of Figure 4.9. This more advanced process allows integration of the filter cell on a pitch of 28 μm, which also corresponds to the minimum realisable CCD pitch. Thus the CCD delay line is here tapped at every cell.

Such a two-level process offers a 4-electrode (two-pair) CCD cell, for 2-phase or 1½-phase operation. Whilst it would seem natural to commonly clock the first electrode pair, and connect the second pair in floating-gate reset mode, using 1½-phase charge transfer, this is not a practicable scheme. A potentially serious problem arises in this technology through the large overlap capacitances between electrodes. Whilst not impairing conventional operation, this overlap can cause severe coupling, or breakthrough, of the clock waveform onto a floating tap. As much as 5V of breakthrough may occur from a 15V clock, reducing the tap potential sufficiently to cut off charge transfer. Resetting the tap to a higher potential is not a viable solution however, since charge transfer out of the tap would then be impeded.

A more elaborate, split-phase clocking scheme has been developed to avoid this clock-breakthrough, though it involves driving each of the four electrodes within the cell separately, as shown in Figure 4.10.

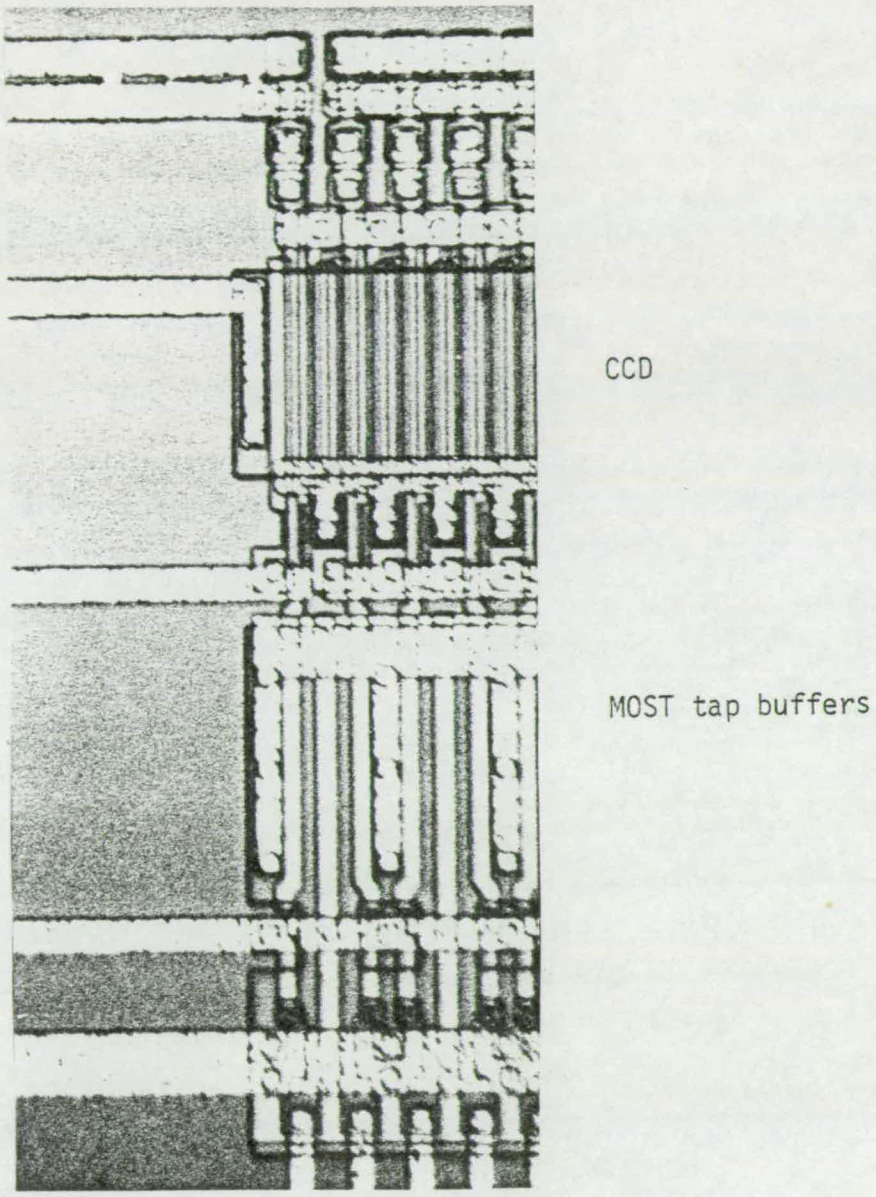


Figure 4.9: Photomicrograph of input stage and initial sections of a polysilicon-gate FGR-tapped CCD delay line

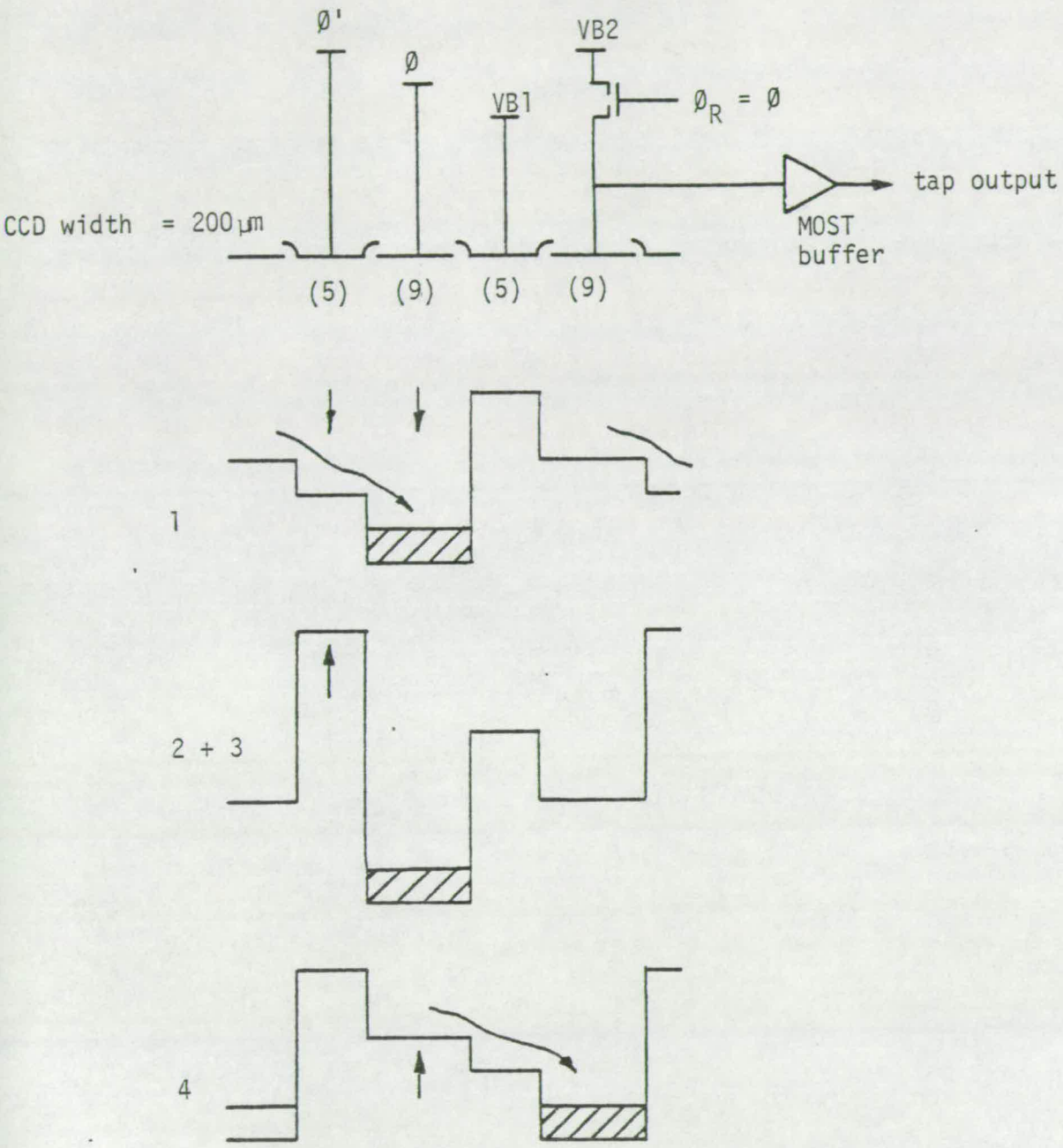


Figure 4.10: Cell schematic of a polysilicon-gate FGR tapped delay line with operating surface potential diagrams

The electrode immediately preceding the tap is clamped to a suitable bias voltage to shield the tap from clock breakthrough in this direction, without impeding the forward transfer of charge. Note however, that the tap must necessarily be followed by a clocked electrode to enable further charge transfer, this being the first electrode of the following cell.

Breakthrough as this clock turns on is not a problem since the tap is normally then reset. (Time frame 1 in the figure). To avoid clock breakthrough onto the tap as this clock turns *off*, the following sequence of operations (corresponding to time frames 2 to 4) is required;

2. Turn off the first electrode in the cell whilst the tap reset is still on, thus avoiding breakthrough;

3. Turn off the reset clock, isolating the tap, and then;

4. Turn off the second cell electrode, transferring charge over the d.c. 'wall', and into the floating tap well.

This action is properly achieved using the split phases $\emptyset$ and $\emptyset'$, shown in the timing diagram of Figure 4.11. Note that phase $\emptyset$ also satisfies the reset clock function, and that steps 3 and 4 occur in convenient sequence as this clock turns off.

A complementary non-overlapping phase, $\overline{\emptyset}$, is required at the input for diode-cut-off operation, and later serves as a tap output sampling clock in a full filter realisation. This modified two-phase clocking scheme, involving only an additional falling edge delay, is much more

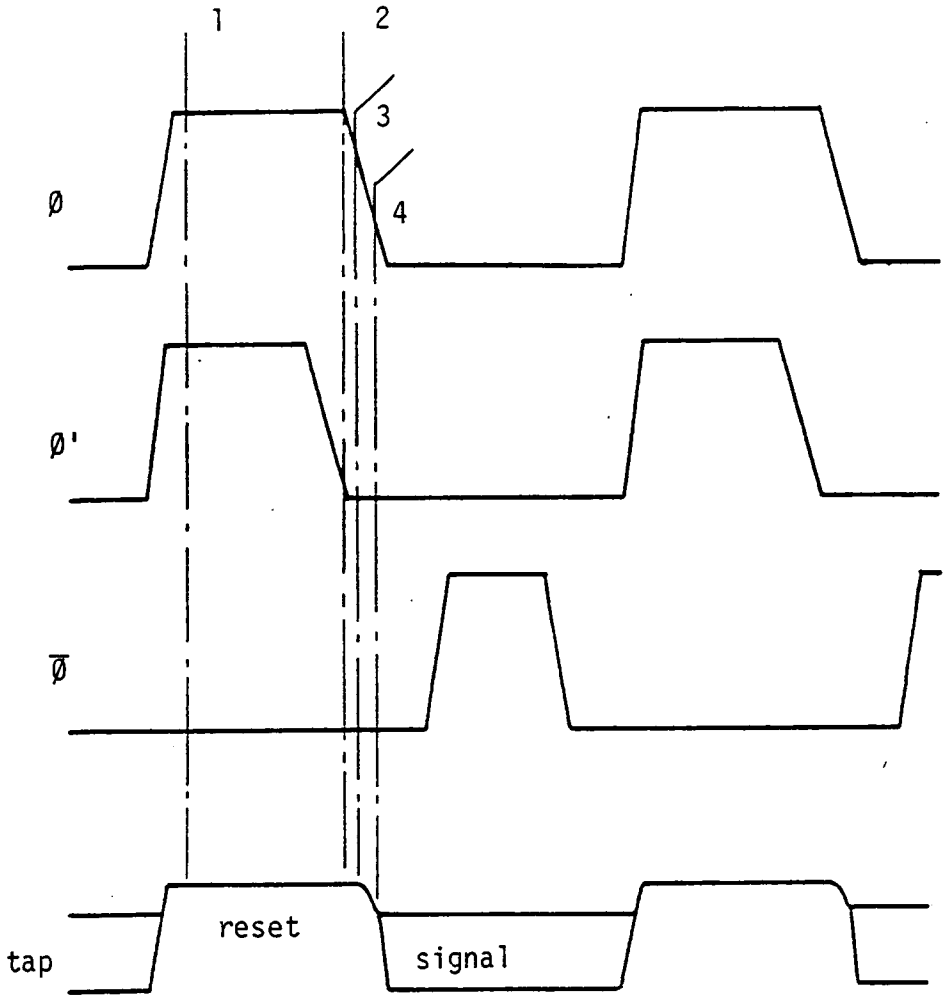


Figure 4.11: CCD and tap output timing diagram

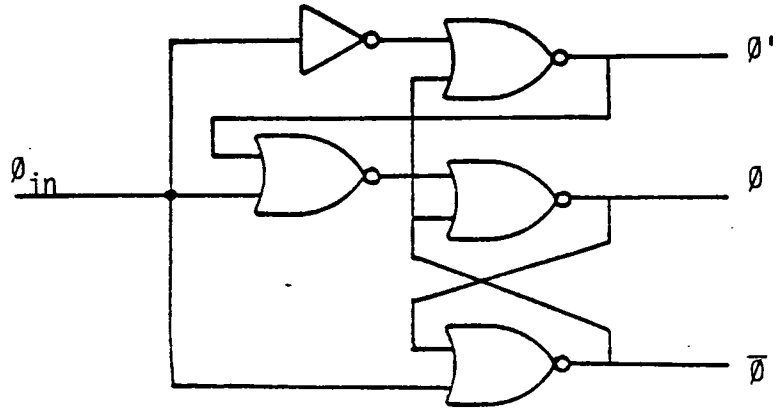


Figure 4.12: Clock generator circuit

easily generated than the overlapping 3-phase scheme. In fact a simple combinational logic circuit is sufficient[65,95] with gate delays ensuring the necessary phase relationships. The clock driver may be implemented in CMOS at 15V to drive the CCD directly, making the delay line easy to use.

The electrode lengths shown in Figure 4.10 comply with the minimum limits set out in the process layout rules. The CCD width (of 200 μm) was chosen to satisfy the $C_0 = C_L$ compromise. In this overlapping gate technology, the tap load capacitance is dominated by overlap of the gates adjacent to the tap, with some contribution from the buffer input. For this configuration;

$$C_0 \simeq C_L \simeq 0.33 \text{ pF}, \tag{4.14}$$

giving rise to the following sensitity relationships;

$$\frac{\partial V_G}{\partial \emptyset_S} = \frac{C_0}{C_0 + C_L} = 0.5 \tag{4.15}$$

$$\frac{\partial Q_S}{\partial V_G} \simeq C_L \simeq 0.33 \text{ pC/V}. \tag{4.16}$$

The input structure of the experimental tapped delay line is shown in Figure 4.13. A dummy tap is included for charge metering, preceded by a sampling gate and input diode for diode-cut-off operation. This scheme has a drawback however. Strong breakthrough of the cut-off gate pulse onto the floating tap occurs through the polysilicon overlap thus disturbing the metered charge just at the moment of cut-off. It is not possible here to protect the tap with a biased blocking electrode, since there must be direct access between diode and tap when the input is on. The experimental TDL was in fact operated succesfully by clamping

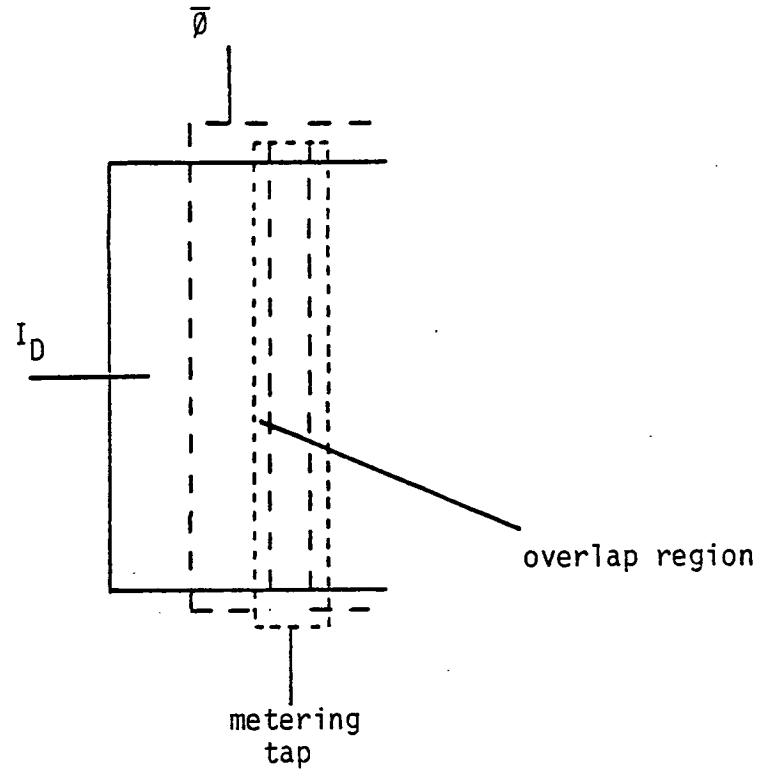


Figure 4.13: Input structure of experimental polysilicon gate FGR tapped delay line

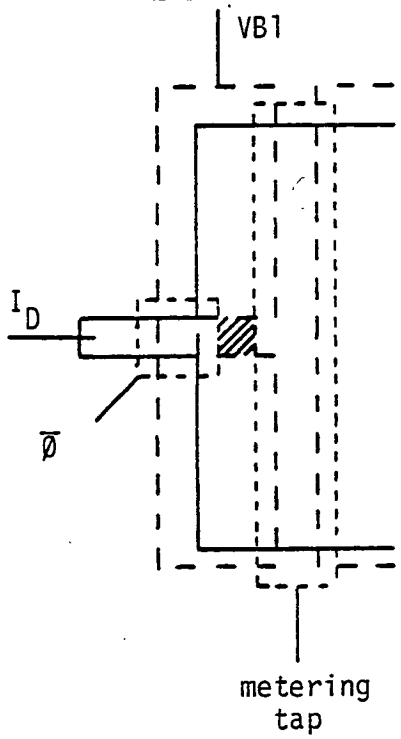


Figure 4.14: Improved input structure

the dummy input tap, but this violates the tap matching condition.

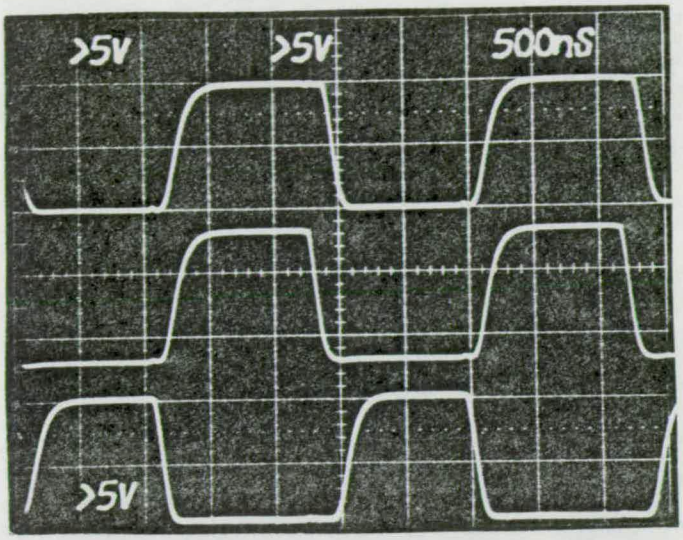
An improved TDL input structure, shown in Figure 4.14, is adopted in the subsequent filter realisation. The input diode and gate are dramatically reduced in area and feed charge to the tap through a small central channel. Over the remainder of the tap width a standard d.c. blocking electrode is included so that, for the most part, this dummy tap appears in an electrical environment identical to the subsequent output taps in the delay line. A small square of floating diffusion (shaded) represents the only mismatch in tap load, and this is readily confined to the order of 1%.

Inherent in this improved structure is a further important advantage. Charge partitioning from under the small cut-off gate is reduced to insignificant proportions. This structure might therefore be employed with advantage in all applications requiring a diode cut-off input technique.

The FGR-tapped polysilicon-gate CCD delay line operated successfully, in discrete form and within an integrated filter. Figure 4.15 shows two photographs that illustrate this performance. The upper photograph depicts clock waveforms conforming to the timing diagram of Figure 4.14, and the lower photograph demonstrates the delay of an input sinusoid at a sample tap output, with the device clocking at 50 KHz. The simple bi-phase reset/signal nature of the tap waveform is evident.

A peak linear tap output signal of 1.5V is attained with 15V clocks before gross distortion effects set in. This is half the value that may be expected from a simple calculation of the range of surface potentials available.

(a)

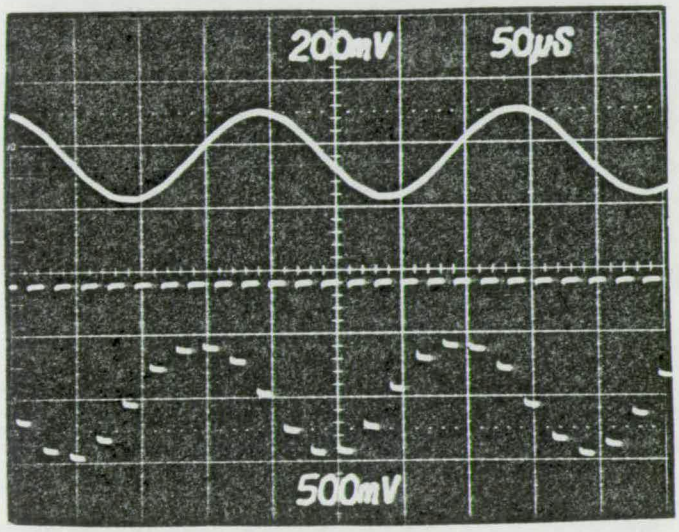


Ø

Ø'

$\overline{Ø}$

(b)



INPUT DIODE

TAP OUTPUT

Figure 4.15: Operation of polysilicon-gate FGR tapped delay line

(a) Clock waveforms
(b) Delay of sinusoid

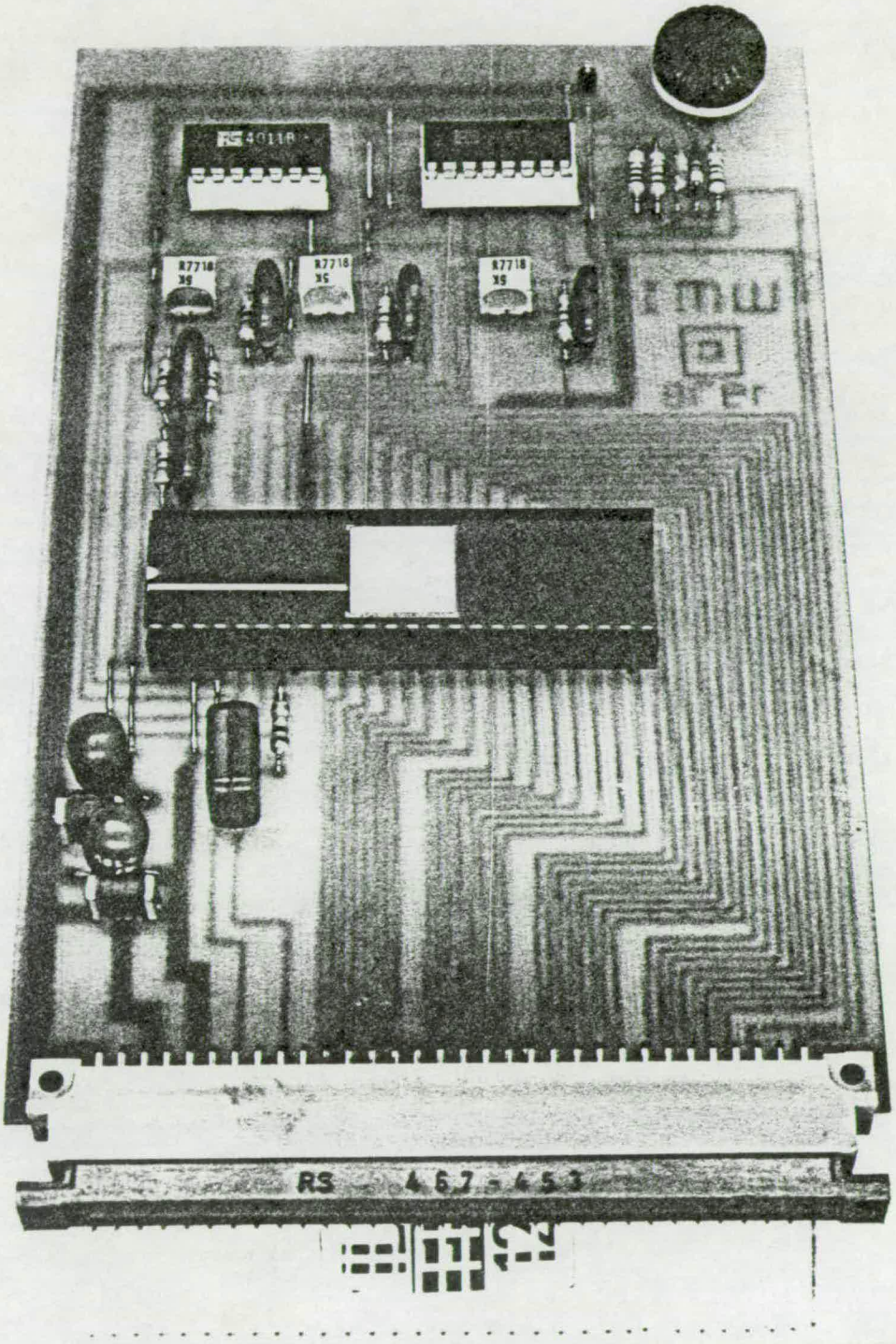


Figure 4.16: Photograph of FGR tapped-delay line module.

Courtesy Wolfson Microelectronics Institute, Edinburgh.

Again the problem of potential 'humps' beneath the interelectrode gaps is a suspected cause; Browne[37] confirms their possible existence on this process. However, because the interelectrode gap is effectively sealed and electrically shielded, the device characteristics remain stable.

Measurements on these devices give c.t.i. at 0.5×10^{-3} per transfer, or effectively 1×10^{-3} per tap, and total harmonic distortion less than -40 dB for short-delay tap output signals at 1v p-p. Cumulative distortion in longer delay lines is considered in the following section.

Although this experimental tapped delay line was not intended for application in its own right, its simplicity of operation has prompted the development of a printed circuit board module, shown in Figure 4.16.† The 32-tap delay line is assembled in a 48-pin package, and is clocked at 15V via two standard CMOS parts, visible at the rear of the board. Spare gates in these devices form an on-board master clock oscillator. The remaining components generate and decouple the necessary bias voltages.

4.5 FEEDBACK LINEARISATION

It is possible to further linearise the CCD transfer function and more significantly, to attain unity through gain, by employing an operational amplifier at the input in a 'feedback-linearisation' scheme.[16,17]

As shown in Figure 4.12, the CCD input stage can be included

---

† I am grateful to J.W. Arthur for this development

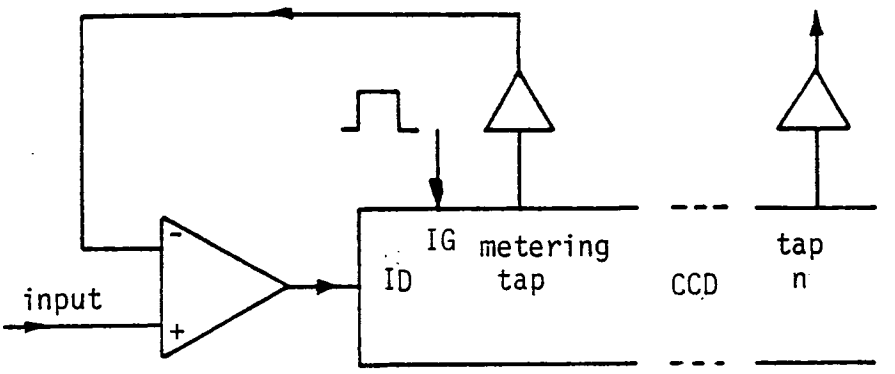


Figure 4.17: Principle of feedback linearisation

within the forward gain path of a voltage-follower connected operational amplifier. The injected charge is thus metered and regulated prior to launching, such that the output from the metering tap is forced to be an exact replica of the input signal , by virtue of the high open loop gain of the operational amplifier. Assuming that the gating pulse allows sufficient time for the closed loop to settle, any harmonic component present in the intrinsic transfer function is reduced by a factor $\frac{1}{1 + AB}$,

where A = Mid-band open loop gain of operational amplifier,

and B = through gain of CCD input stage.

Furthermore, the overall transfer function between the signal input and any tap output now has a gain;

$$A_V = 1 - \frac{1}{AB} , \qquad (4.17)$$

which may be very close to unity.

The power of this technique should not be underestimated, since it achieves a linear, lossless transfer function by automatic compensation; its effectiveness is limited only by tap matching within individual chips. Other techniques rely heavily on accurate anticipation and control of processing parameters, such as oxide thickness and doping density.[94]

An operational amplifier to perform this task of feedback linearisation has been successfully developed; details are reported in Chapter 6.

The nearly perfect CCD transfer function that is achieved with this technique is illustrated later in this Thesis in Figure 8.7. This shows a chirp waveform delayed through a 256 tap CCD with unity gain maintained to within 1%.

4.6 EFFECTS OF TAP MATCHING AND CHARGE TRANSFER INEFFICIENCY ON THE CCD TRANSFER FUNCTION

The tapping schemes discussed here depend for effective operation upon good matching. Conditions at the input tap must be duplicated at all of the output taps if identical replicas of the input signal are to be produced. In practice the taps will not be identically matched and therefore some error in gain, and in linearity may be anticipated.

Consider the charge-voltage sensitivity function, $T = V_{tap}/Q_S$, given in (4.1). The first order sensitivities of T to variations in each of the components C_0, C_L and C_D are

$$\frac{\frac{\partial T}{T}}{\frac{\partial C_L}{C_L}} = \frac{\partial T}{\partial C_L}\cdot\frac{C_L}{T} = -\frac{C_L\,(C_D + C_0)}{C_0C_D + C_DC_L + C_0C_L} \simeq -0.91 \quad (4.18)$$

$$\frac{\frac{\partial T}{T}}{\frac{\partial C_0}{C_0}} = \frac{\partial T}{\partial C_0}\cdot\frac{C_0}{T} = 1 - \frac{C_0(C_D + C_L)}{C_0C_D + C_DC_L + C_0C_L} \simeq 0.09 \quad (4.19)$$

$$\frac{\frac{\partial T}{T}}{\frac{\partial C_D}{C_D}} = \frac{\partial T}{\partial C_D}\cdot\frac{C_D}{T} = -\frac{C_D(C_L + C_0)}{C_0C_D + C_DC_L + C_0C_L} \simeq -0.16 \quad (4.20)$$

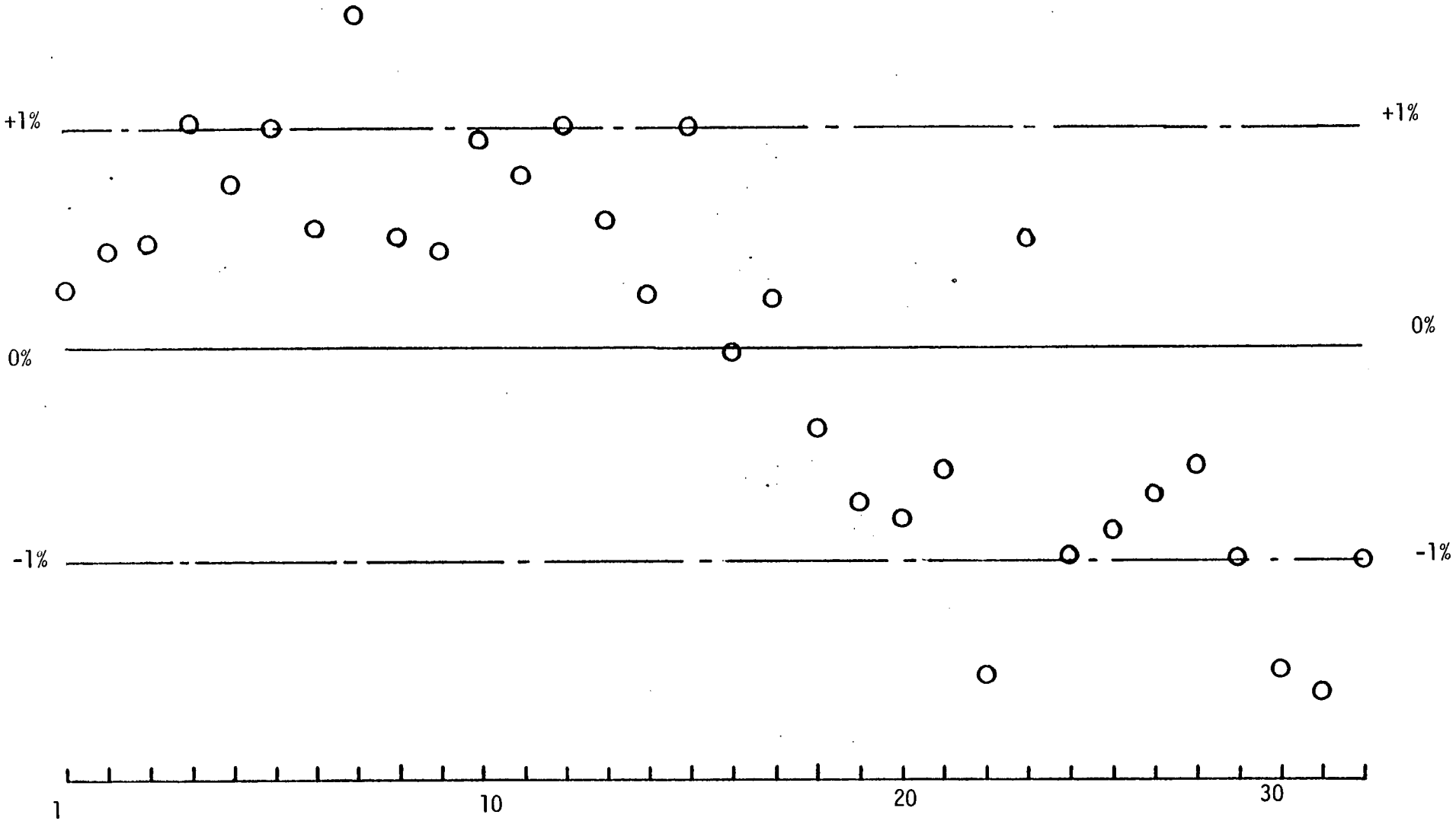


Figure 4.18: Measured tap gain deviation along an operational 32 tap delay line

To first order then the tap gains are most sensitive to variations in the load capacitance C_L (normally predominantly thin oxide), and to a lesser extent to variations in channel depletion capacitance C_D (controlled by substrate doping density). Figure 4.18 shows measured variations of in the gain of 32 taps along one of the experimental tapped delay lines. There is a resultant r.m.s. tap gain error of 1%.

Clearly the relevant effect of small parametric variation is confined to gain errors only. The linearity of the transfer function T, is good to within 1%, so that parametric variations of 1% will introduce variations in linearity of some 0.01%; which are insignificant.

A more significant effect on linearity may be that of signal distortion by non-linear charge transfer inefficiency, ε. It is commonly assumed that ε is a constant, based on expected linear fringing fields at the edgges of each transfer site, as shown in Figure 4.19(a). Charge packets of increasing size are exposed to a linearly increasing area of fast surface states, giving rise to a constant value of ε. The effect on charge packets is one of frequency-dependent attenuation guaged by the cumulative c.t.i., $N\varepsilon$, as shown in Figure 3.8. For constant ε no harmonic distortion occurs.[8] The linear field assumption is generally false, but often a region within the potential wells may be found which sufficiently approximates.

Small amounts of transfer function distortion may be expected with the input and tapping schemes proposed here, since the charge packets are not perfectly linear representations of the signal samples. Thus although the charge packets may be linearly attenuated, the

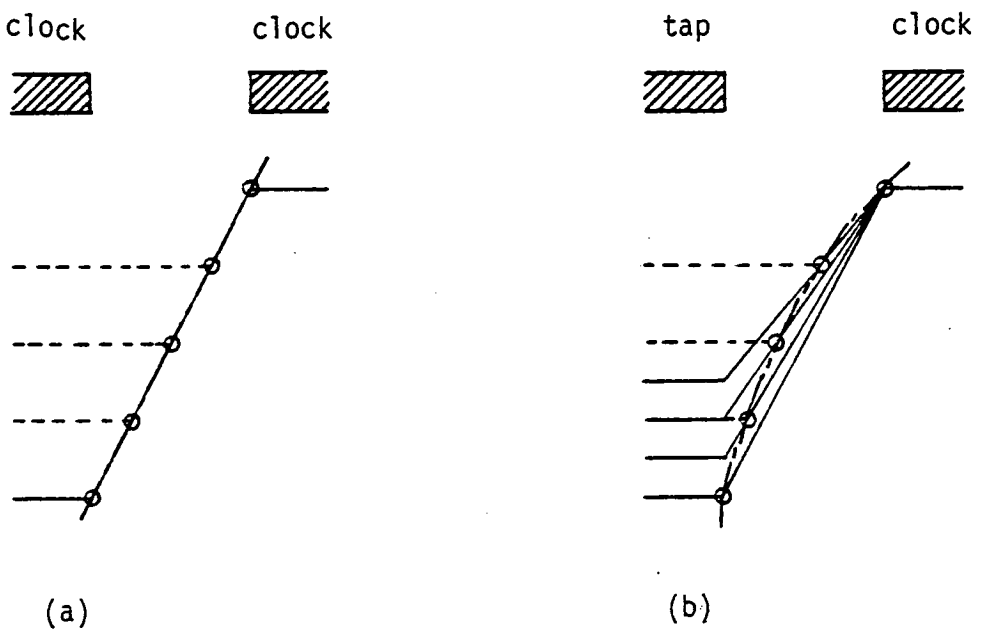


Figure 4.19: Fringing field effects on exposed surface area loci

(a) Conventional clocked electrodes

(b) Floating gate tap

signal information will become distorted. More significantly, the constant ε approximation does not hold for an FGR tap.

Consider the FGR tap surface potential diagram shown in Figure 4.19(b). As charge is injected, the depth of the well is reduced through variation of the electrode voltage. This causes charge packets of increasing magnitude to be exposed to a non-linearly increasing area of fast surface states, as shown by the exaggerated locus. A constant ε no longer applies, and direct cumulative distortion of charge packets occurs. We may expect this distortion to vary with signal frequency as well as with the number of taps passed. Figure 4.20 shows typical measured harmonic components in a 1V p-p signal as a function of relative signal frequency after transfer through 256 taps, the signal having been feedback-linearised at the input.

Note that there is a definite, complicated frequency dependence. Since this is a cumulative effect, shorter delay lines demonstrate a correspondingly improved performance.

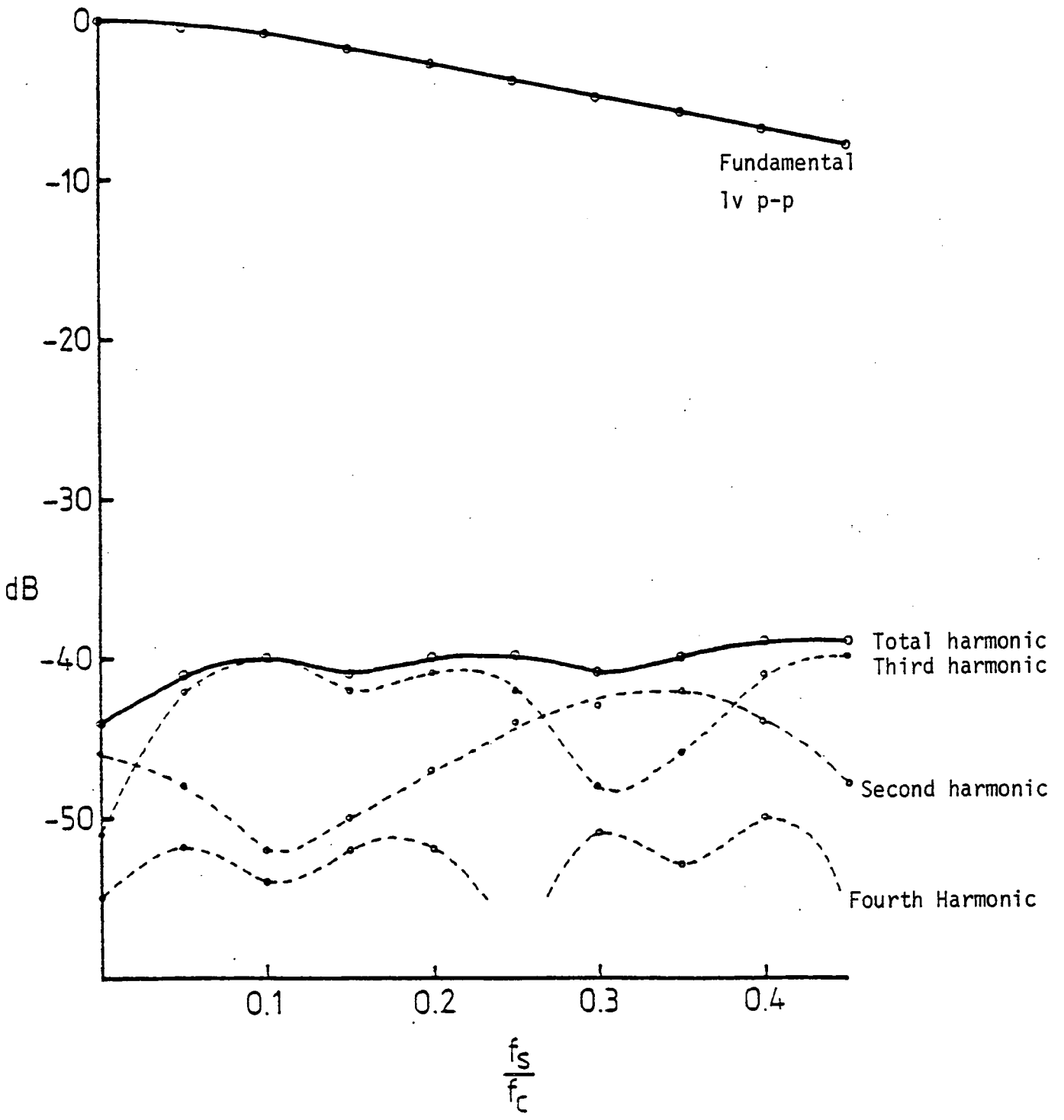


Figure 4.20: Frequency dependent signal attenuation and distortion after 256 FGR taps

CHAPTER FIVE : MULTIPLICATION TECHNIQUES

If a programmable transversal filter containing large numbers of points is to be successfully integrated, then a compact, accurate voltage multiplier will be required at each point. Furthermore, the outputs of all multipliers must be summed on-chip to provide a single filter output. These features are elegantly provided by the MOST transconductance multiplier, which forms the product of two voltages into a current output that can be commonly summed within an array of multipliers. These techniques have previously been reported by Radeka et al[26,27], who considered the two-transistor 4-quadrant multiplier. This chapter reports improvements to these techniques that realise one-transistor 4-quadrant and 2-quadrant multipliers.

5.1 FUNDAMENTAL THEORY

Consider a first order approximation to the drain current flowing in an MOST in the triode region:

$$I_S = \beta[(V_{GS} - V_T)V_{DS} - V_{DS}/2] \tag{5.1}$$

where V_{GS} and V_{DS} are defined in Figure 5.1. Note that a possible product term, $V_{DS}V_{GS}$, exists. If V_{DS} is small then;

$$I_S \simeq \beta(V_{GS} - V_T)V_{DS}\,. \tag{5.2}$$

This approximation forms the basis of the well-known voltage controlled resistor[4] and has been exploited by Lampe[50] and others in programmable

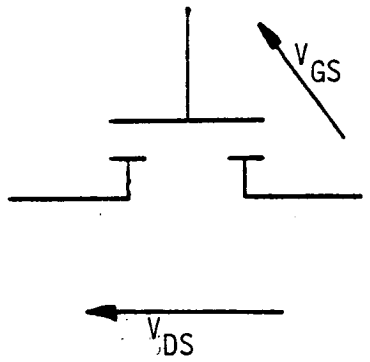


Figure 5.1: MOS transistor definitions

filter realisations. However the restrictions applying to the approximation enforce a poor trade-off between linearity and dynamic range. A more suitable multplier may be obtained by properly cancelling the undesirable terms in (5.1).

The principle of cancellation requires the formation of a reference current, I_0, which contains all of the unwanted terms. Thus we require to form:

$$I_0 = \beta[(V_{GS_o} - V_T)V_{DS} - V_{DS}/2] \quad (5.3)$$

where V_{GS_o} is a gate offset required to bias the MOST well into the triode region, where these equations apply.

The product current I_P may then be found by subtracting the reference current I_0 from I_S:

$$I_P = I_S - I_0 = \beta(V_{GS} - V_{GS_o})V_{DS} \quad (5.4)$$

or

$$I_p = \beta V_x V_y \quad (5.5)$$

where $V_x = V_{GS} - V_{GS_o}$, and $V_y = V_{DS}$.

Note that, for suitably large V_{GS_o}, signals V_x of either sign are correctly accommodated. Signals V_y which are negative interchange the definition of the source and drain terminals, but (5.4) still applies since it merely contains an expression of a *difference* in gate voltage and does not depend upon any absolute value relative to the source. Thus 4-quadrant multiplication is possible in principle, by applying V_x with respect to a reference gate voltage V_{GS_o}, and V_y across the diffused (source and drain) terminals.

Two realisations of the 4-quadrant multiplier algorithm are shown in Figure 5.2(a) and (b). Both employ external summing amplifiers which are common to a full array of multipliers. These amplifiers should provide the necessary feature of a virtual earth along the summing bus to which the applied voltages V_y are referred, and may transduce the currents to more useful voltages.

The conventional 4-quadrant multiplier shown in Figure 5.2(a) uses two transistors to form separately but simultaneously, the two current components I_S and I_0. These are summed individually and the output is then formed from their difference. This technique suffers two disadvantages. Firstly the transistor pair must have carefully matched gains and thresholds. In practice this matching will never be perfect, and a small difference in gain between the transistors, of factor $\Delta\beta/\beta$, results in the following product error;

$$I_P' - I_P = \frac{\Delta\beta}{\beta}\left[(V_{GS} - V_T)V_{DS} - V_{DS}^2/2\right] \tag{5.6}$$

where I_p' is the product in error and I_p is the ideal product. Similarly, a difference in threshold voltage of ΔV_T between the transistors results in a product error;

$$I_P' - I_P = \beta[\Delta V_T V_{DS}]. \tag{5.7}$$

The second disadvantage of this technique stems from the requirement of laying out the two matched transistors in close proximity. This may prove to be impossible within the small cell pitches available in a high density filter realisation.

The time-multiplexed single-transistor 4-quadrant multplier of Figure 5.2(b) was developed to overcome the disadvantages inherent in

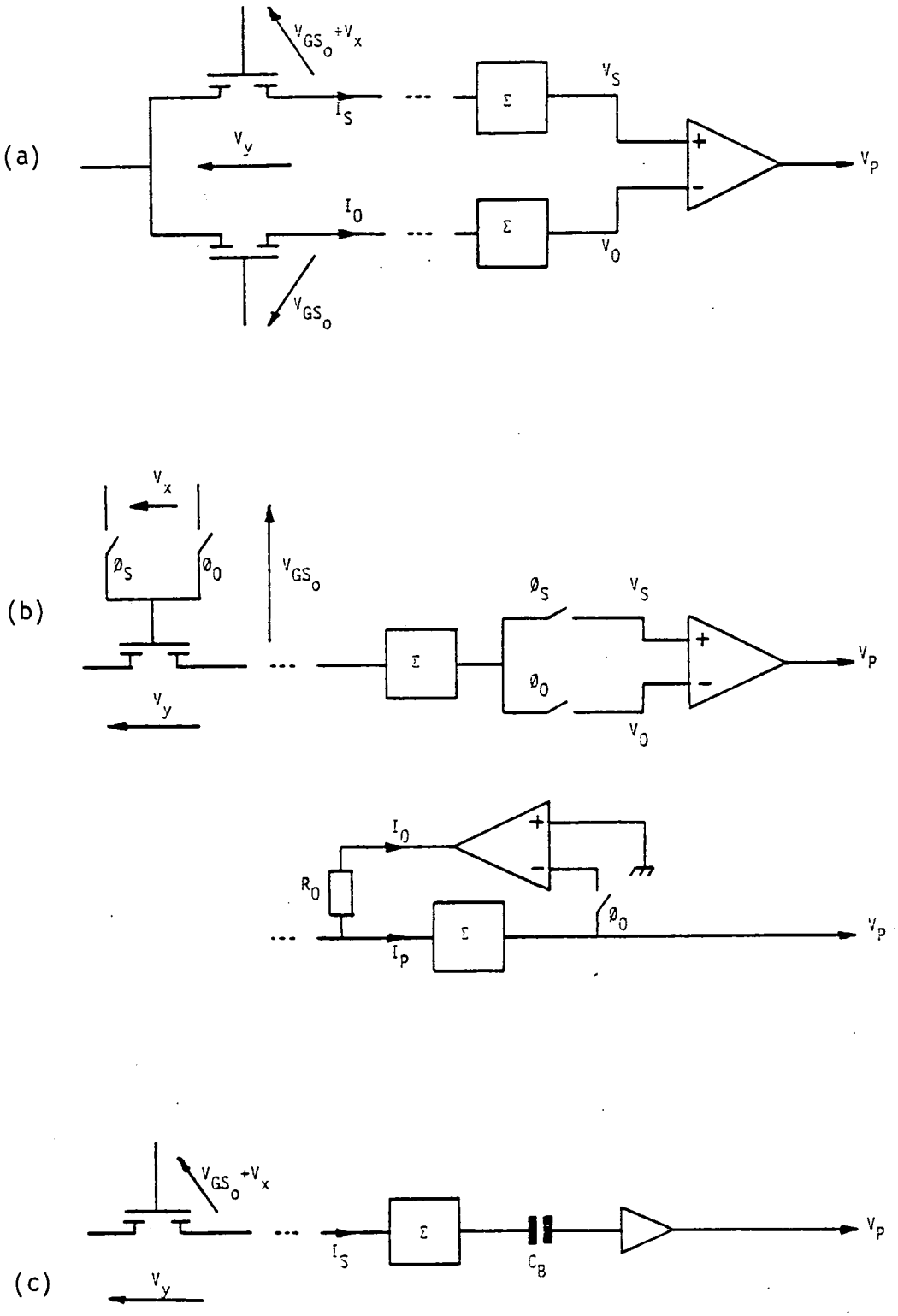


Figure 5.2: MOST multiplication schemes

(a) 2-transistor, 4-quadrant
(b) 1-transistor, 4-quadrant
(c) 1-transistor, 2-quadrant

the conventional two-transistor version. The two current components, I_S and I_0, are here formed through the same channel during two consecutive time slots by alternately switching V_{GS}, and then $V_{GS_0} + V_x$, onto the gate. The output components V_S and V_0 are separately held and then differenced to form the required product V_P. A variation on this technique is also shown in the figure, whereby I_0 is actually cancelled from the summing bus using a bleed resistor R_0. During the I_S phase, I_0 remains cancelled, leaving only the desired product I_P to occupy the full dynamic range of the summing amplifier.

Because only one channel is used, all matching errors are excluded, making this a more accurate technique. Furthermore only a single multiplying transistor must be included at each filter point. Although the time-multiplexed operation relies on a clocking scheme, this fits naturally into any CCD application. Figure 5.3 demonstrates correct 4-quadrant multiplication of two sinusoids at 3 kHz and 300 Hz using the time-multiplexed single-transistor technique.

Further simplification of the multiplier is possible in applications where multiplication by a *constant* weighting coefficient is required. If this weighting coefficient is applied as V_y, then the information given in the reference current I_0 is static, and forms a quiescent component at the output which is unwanted in any a.c. coupled system. Since a primary application of the PTF is as a stationary-weighted a.c. coupled matched filter, this reference may be entirely removed, leaving the simplified single-transistor multiplier of Figure 5.2(c). Tap signals are applied as changes in the gate voltage, ΔV_{GS}, on top of the biasing level, V_{GS_0}, and, from (5.1)

$$\Delta I = \beta \Delta V_{GS} V_{DS}, \text{ or}$$

$$\Delta I = \beta V_x V_y \tag{5.8}$$

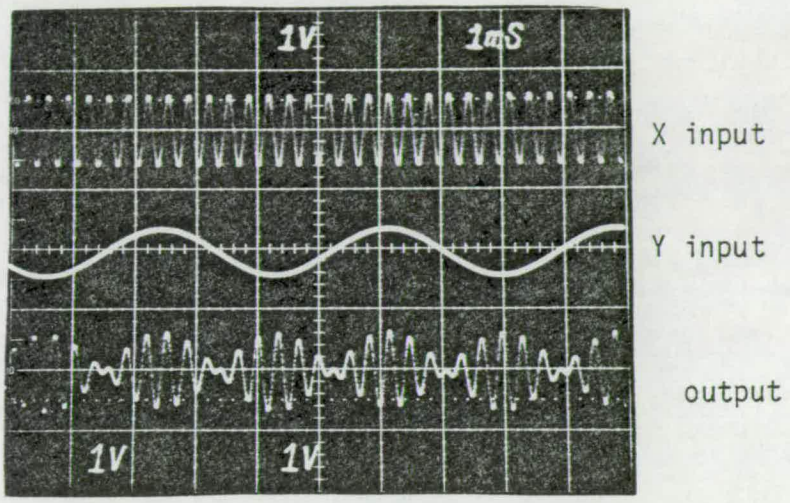


Figure 5.3: 4-quadrant multiplication of two sinusoids using the time-multiplexed single-MOST multiplier

f_x = 3kHz, f_y = 300 Hz

After summation these changes are detected through the quiescent blocking capacitor C_B.

5.2 SUMMING TECHNIQUES

The summing amplifiers shown in Figure 5.2 are required to linearly transduce a current I, summed on a single bus, into an output voltage V, whilst preserving a virtual reference, V_{y_o}, on the summing bus. The two circuits shown in Figure 5.4. achieve this specification. The first circuit employs an operational amplifier and feedback summing resistor, R_B. Through negative feedback the summing node is maintained at the desired reference potential, whilst the current on the bus is deflected entirely through R_S, across which the desired output is realised. This summing circuit can be slow to settle however, limiting its use in general to applications below approximately 100 kHz.

The common-base bipolar stage offers high speed operation in a simple circuit, and is therefore preferred to the operational amplifier scheme. Here a reference voltage is applied to the pnp transistor base which thus provides a virtual reference voltage at the emitter, through the nearly constant V_{BE} drop achieved in the active region. Current, I, injected at this node passes through the transistor to the collector, and thus through the summing resistor R_S. Resistor R_B is provided to bias the stage into the active region. R_B and R_S must be chosen such that correct operation is maintained over the full expected range of multiplier currents.

This stage may not be integrated with the remainder of the filter, however by virtue of its simplicity this is not considered to be any disadvantage.

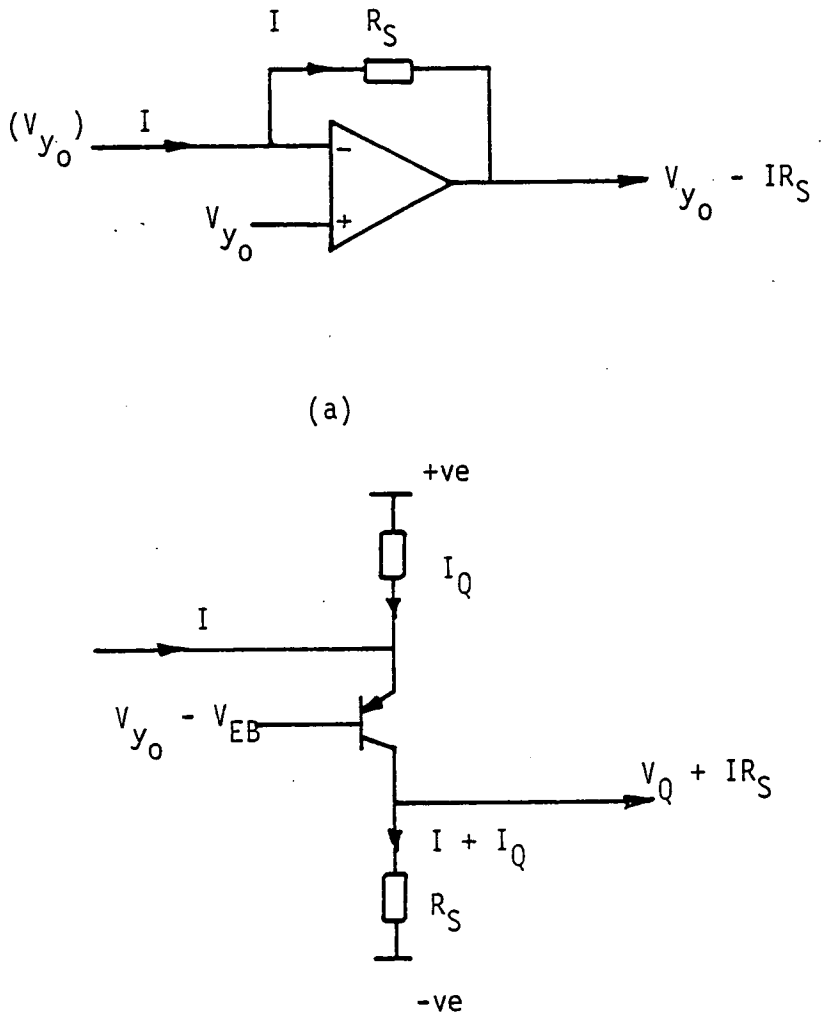


Figure 5.4: Summing amplifier circuits

(a) Operational amplifier stage
(b) Common base bipolar stage

5.3 SECOND ORDER LIMITATIONS

Consideration of first-order equations has shown the MOST transconductance multiplier to be theoretically perfect. It is possible that second-order effects may limit its performance. Of the three second-order terms considered in section 3.10 that of saturated channel length modulation does not apply here, and threshold voltage modulation has no effect, since all threshold terms are cancelled in the multiplication algorithm. The remaining effect of field-dependent mobility is relevant however and becomes the dominant source of distortion within the multiplier.

Field-dependent mobility causes the gain, β, of the MOST to vary with the applied gate voltage. Figure 5.5 shows the effect of this term upon the transconductance, g_m, of the MOST, as a function of applied gate bias, using the field-dependent mobility expression of (3.15). The ideally constant value of transconductance, for gate bias into the triode region, assumes a hyperbolic function. Measurements of the transconductance of the multipliers used in the first of two PTF designs are shown in Figure 5.7. They clearly demonstrate this gate modulation effect. Note that operation at gate drives deep into the triode region should give improved linearity through a more constant transconductance curve.

Consider the modified expression for drain current at gate bias V_{GS_o};

$$I_o = \frac{\beta[V_{GS_o} - V_T)V_{DS} - V_{DS}^2/2]}{1 + \theta(V_{GS_o} - V_T)} \qquad (5.9)$$

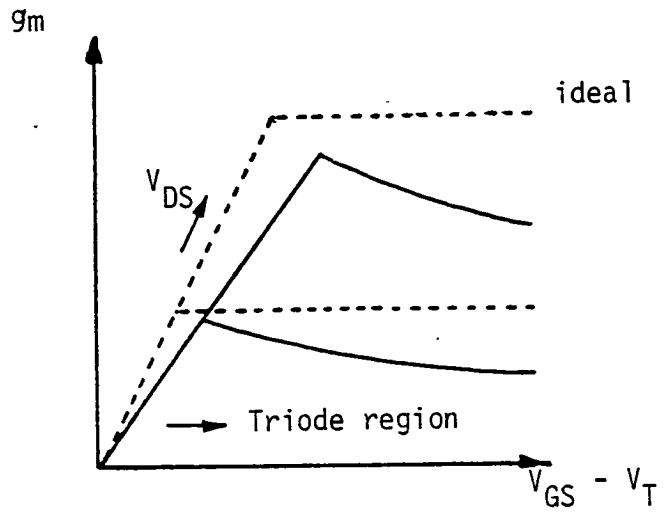


Figure 5.5: Effect of field-dependent mobility upon MOST transdconductance

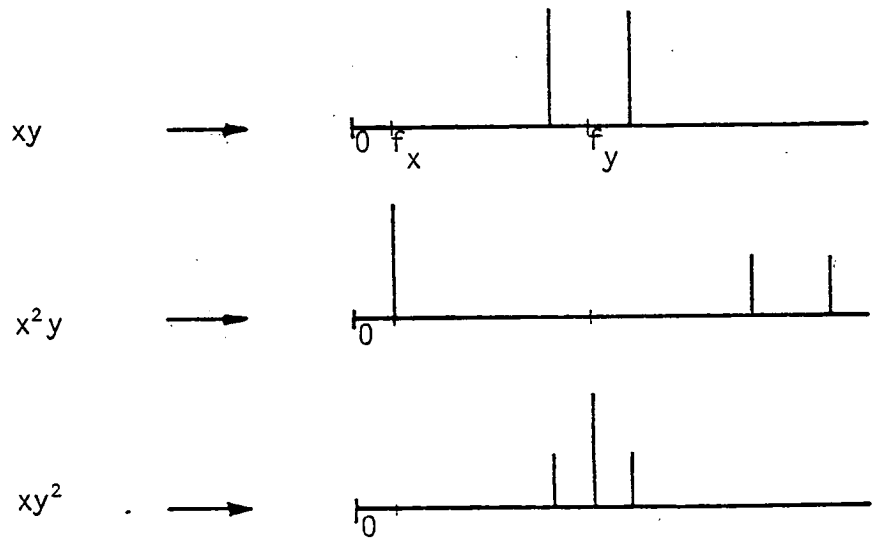


Figure 5.6: Effect of field dependent mobility on spectral distribution of product sinusoids

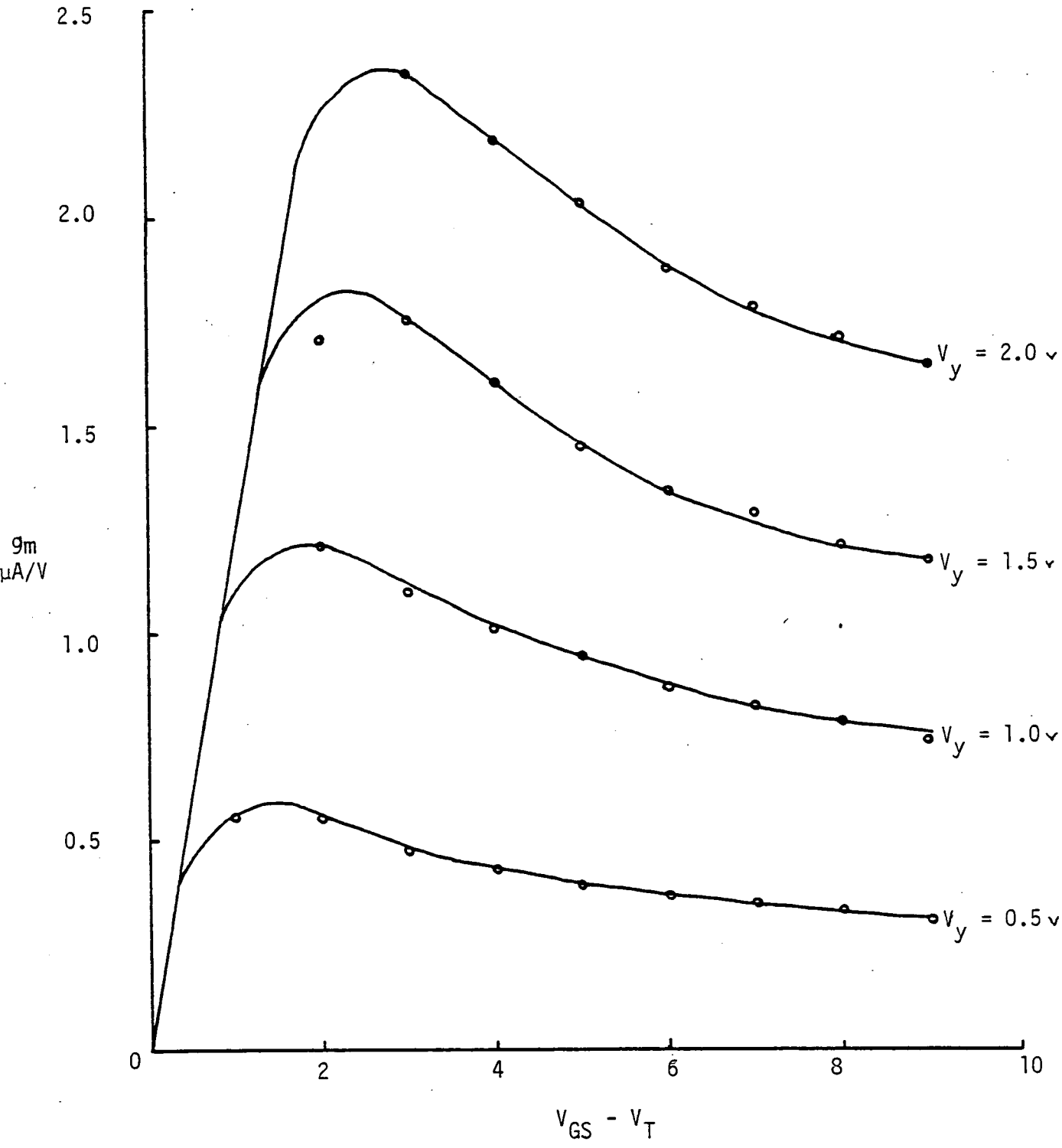


Figure 5.7: Measured transconductance characteristics for an MOST multiplier

Now apply a signal, V_x, to the gate, giving an expected change in current I_p;

$$I_p = \frac{\beta[(V_{GS_o} - V_T)V_{DS} - V_{DS}^2/2] + \beta V_x V_{DS}}{[1 + \theta(V_{GS_o} - V_T)][1 + kV_x]} - I_0 \qquad (5.10)$$

where $k = \theta[1 + \theta(V_{GS_o} - V_T)]^{-1}$; which is $<<1$ and reduces for increasing triode region drive, $(V_{GS_o} - V_T)$. Taylor expansion of the second term in the denominator and rationalisation of significant terms gives;

$$I_p \simeq \beta'(V_x V_{DS} - kV_x^2 V_{DS} + \frac{k V_x V_{DS}^2}{2} + \ldots) \qquad (5.11)$$

where all further terms are of order k^2 or higher, and may be neglected, and the modified gain constant β' is given by:

$$\beta' = \frac{1 - k(V_{GS_o} - V_T)}{1 + \theta(V_{GS_o} - V_T)} . \qquad (5.12)$$

There is clearly some attenuation of the product, as may be expected, coupled with distortion of both signal, V_x, and reference, $V_y = V_{DS}$, values. The theoretical effect of these terms upon spectra relevant to the multiplication of sinusoids is shown in Figure 5.6. Note that distinct patterns of peaks are formed supplementary to the expected sum and difference frequency components. Figure 5.8 demonstrates these effects exactly in the measured spectrum of the product of sinusoids shown in Figure 5.3. This 4-quadrant multiplier is typically capable of linear multiplication with all harmonic products below -36 dB.

Where a simplified multiplier for stationary tap-weights is used, the reference distortion merely represents a source of (coherent) tap-

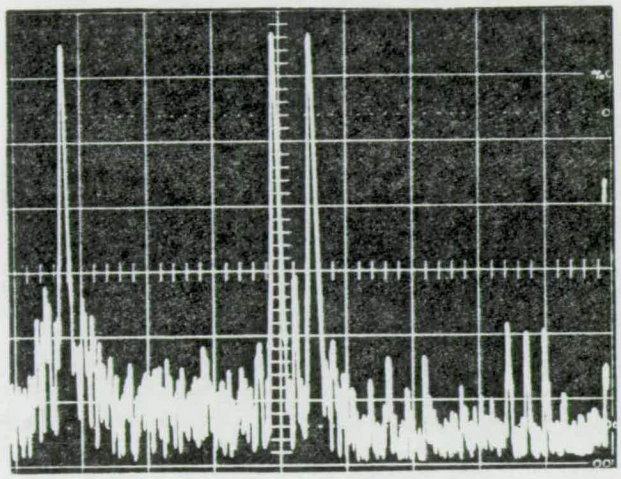


Figure 5.8: Measured spectral distribution of product of sinusoids for an MOST multiplier

weight error and only the applied signal suffers harmonic distortion, through (5.11). Figure (5.9) shows the output spectrum of a typical stationary weighted sinusoid, using a deep depletion MOST multiplier. Note that operation deeper into the triode region here allows linear weighting to -46 dB t.h.d. using typical signal magnitude and bias conditions.

5.4 MATCHING ERRORS

As a multiplier is required at each point in these filter realisations, consideration must be given to the effects of parametric mismatching upon multiplier uniformity. Threshold variation may be expected, but no threshold related terms appear in the product algorithm, so this parameter may be ignored. Variations in the transconductance of the multiplier appear as a direct tap weight error however. The dominant effect here is likely to be oxide thickness variation, typically of the order of 1% across a chip. It is interesting to note that split-gate filters also suffer from this effect.

5.5 DRIVING LIMITATIONS

If these multipliers are to be used within a monolithic filter, then the signal and reference voltages will be supplied from on-chip buffers, and their effect on multiplier performance and matching must be considered. Source follower buffers with constant current loads exhibit small signal gain variations typically less than .1% and are thus ideal for this application. Threshold and gain variations can lead to d.c. offsets around $\pm$10 mV however. Such offsets in the *signal* channel of the multiplier are again nicely cancelled in the product algorithm. Offset variations in the *reference* channel add directly to the applied tap weight value, and thus we may expect a random tap weight error of the order of $\frac{1}{2}$% from this effect in systems using a tap weight range of $\pm$1V.

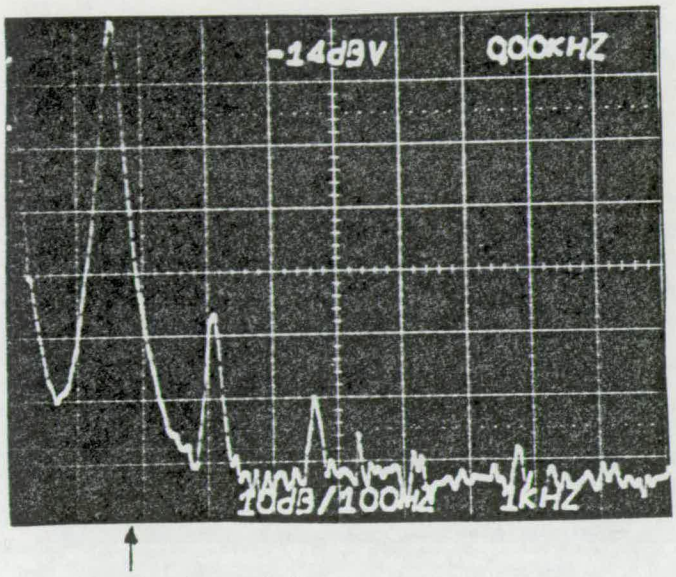


Figure 5.9: Output spectrum of stationary weighted sinusoid, obtained from deep-depletion MOST multiplier

1.5 kHz, 1v p-p

The reference buffer must also supply the multiplier channel current from a finite output conductance. This will cause a predictable and harmless attenuation, but may also lead to non-linear tap weight distortion, and to harmonic distortion of the weighted signal. Consider the tap weight buffer and multiplier configuration of Figure 5.10. A pre-buffer tap weight V_y' is applied to achieve a real tap weight V_y at the multiplier. In general, the buffer load current I_L is limited by a power budget, and β_A is made large to achieve a good output conductance. Similarly β_M is made small to minimise the loading effect. Typically β_A will be 100 x β_M. Despite this, the quiescent multiplier current I_0 may easily be of the same order as I_L, and will thus affect the buffer transfer function. Analysis of this effect is complicated, but fortunately under these conditions the error caused is marginal, even up to the limiting point where $I_0 = I_L$, and the active buffer transistor cuts off. Shown in Figure (5.10) is a SPICE simulation of the transfer function of a tap weight buffer with and without a multiplier load. Above the saturation point mentioned, distortion of weight values up to $\pm$1V remains at less than $\frac{1}{2}$%.

If a small signal, v_x, is now applied to the multiplier gate a product current, i_p, will be stimulated;

$$i_p = \frac{\beta_M v_x v_y}{1 + r_o \beta_M v_x}, \qquad (5.13)$$

where r_o = buffer output resistance (small signal).
Clearly this differs from the ideal case through the non-unity denominator. Now $\beta_M r_o \ll 1$ so,

$$i_p \simeq \beta_M v_x v_y (1 - r_o \beta_M v_x + ...), \qquad (5.14)$$

giving a dominant second harmonic signal distortion at 20 $\log_{10} r_o \beta_M$.

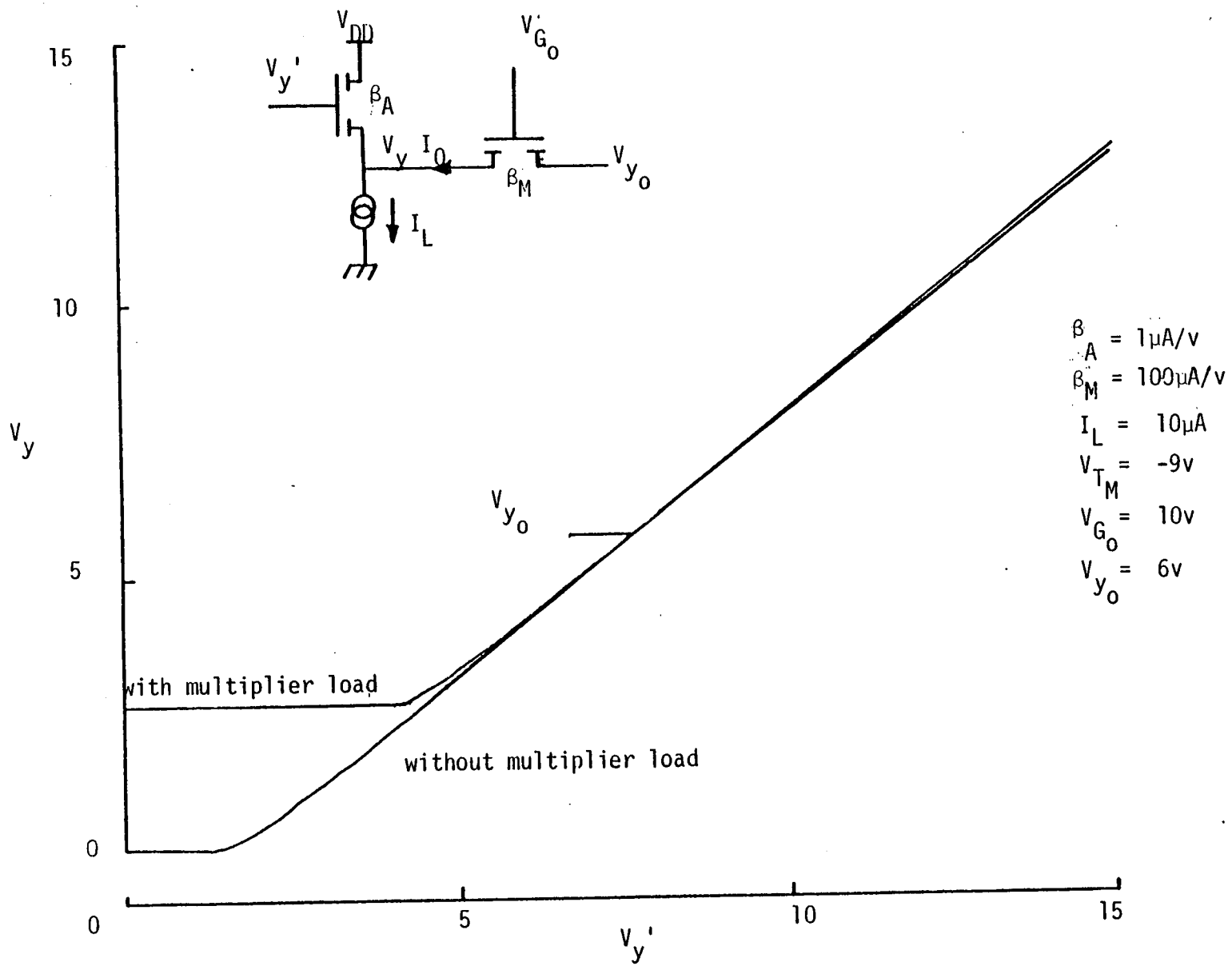


Figure 5.10: SPICE simulation of effect of multiplier load on tap weight buffer transfer function

In the case of the buffer/multiplier under consideration,

$$r_o \simeq 10k\Omega,$$

and

$$\beta_M \simeq 0.5\ 10^{-6}$$

$$\therefore\ 20 \log_{10} r_o \beta_M = \underline{-46\ dB} \qquad (5.15)$$

Just such distortion is visible in the measurement shown in Figure (5.9), and we may deduce that the intrinsic distortion, caused by mobility dependence, is in this case so low as to be masked by a dominant effect induced by a finite driving impedance.

5.6 SUMMARY

The MOST transconductance multiplier has been shown to be a simple, useful device for high-density parallel multiply-add structures. The single transistor form offers advantages in both packing-density and accuracy. Consideration of the intrinsic properties of the device, and of the effect of compatible on-chip drivers, has shown that tap weight accuracy and signal linearity of the order of 1% may be achieved under normal operating conditions.

CHAPTER SIX : AN OPERATIONAL AMPLIFIER

The requirement for a monolithic CCD-compatible operational amplifier has already been identified in section 4.5 on feedback linearisation. Further applications of a suitable design, within these and other circuits, are manyfold. Such an operational amplifier must be simple, compact, low-power, and for the broadest possible range of application offer a wide operating bandwidth. Due consideration must be given to internal compensation where necessary, though this may well increase the layout area. Furthermore, correct biasing of the amplifier should be achieved internally.

6.1 CLASSICAL APPROACH

The monolithic single channel MOST operational amplifier has attracted the attention of several workers, notably Fry[76], Weste[77], Tsividis[78], and Senderowitcz.[79] Of these, the latter two have refined the design to that of a generally useful 'black-box'. All of the designs reported correspond to the classical architecture shown in Figure 6.1, and adopt circuit realisations which correspond to common bipolar counterparts.

The simplest possible single-channel MOST amplifier that can be designed according to these principles is shown in Figure 6.2. Note that such a realisation will contain at least thirteen MOS transistors nine of which, being in the differential and gain stages, demand exacting design. Indeed design of the differential stage is extremely critical, and in operation the amplifier will be most sensitive to the tolerance of parameters within this stage. Furthermore, the complicated pattern

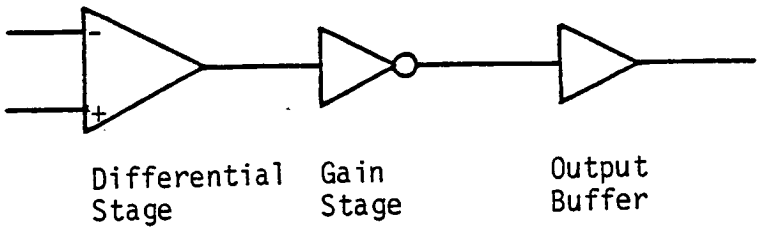


Figure 6.1: Operational Amplifier Architecture

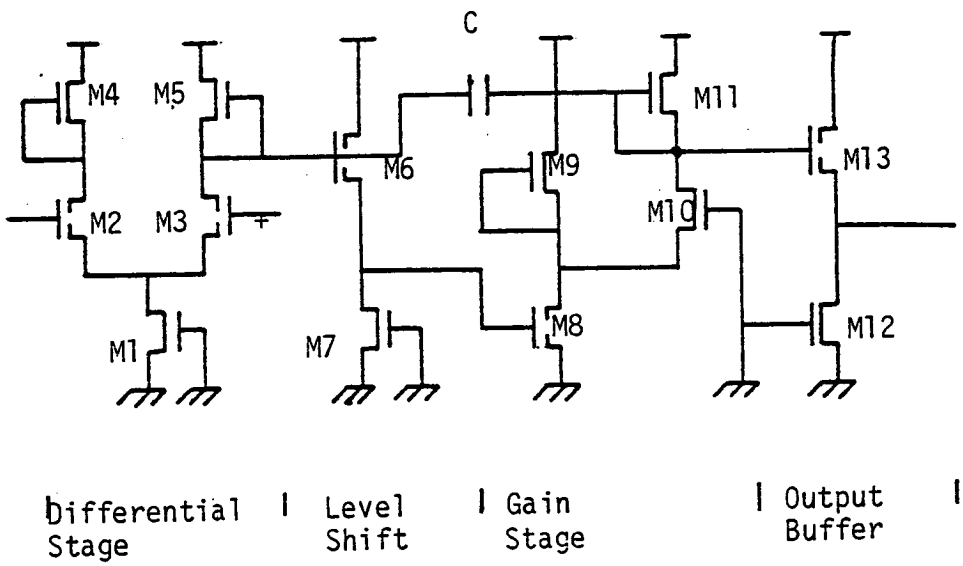


Figure 6.2: Typical MOST operational amplifier circuit

of poles which results in the open loop response often necessitates the incorporation of a large compensating capacitor, C, of the order of 5pF or more. The area occupied by such a design is typically 60,000 μm^2.

6.2 SWITCHED CAPACITOR DIFFERENTIAL STAGE

For applications associated with clocked circuitry, such as the CCD, a much simpler, insensitive differential stage can be provided by using the switched capacitor approach shown in Figure 6.3. With switches $\emptyset$ closed, node A is set to V_+, and node B is reset to some reference potential. As switches $\emptyset$ open and $\overline{\emptyset}$ close, node B is allowed to float, and node A undergoes a change in potential of $V_+ - V_-$. This change is transmitted through capacitor C to node B. If any stray capacitance exists at node B then the differential signal is attenuated, but remains a linear representation of the difference in voltage between the two inputs.

The reset level at node B may be chosen to suit any following high gain stage within the amplifier, obviating any level shifting operation. Minimum geometry MOSTs will serve as switches, and for monolithic realisations values of C around 1pF suffice to dominate stray effects.

The switched capacitor differential stage has no critical design parameters, offers a common-mode range comparable to the clock amplitude, has perfect common-mode rejection and requires no preceding level adjustment. It does not offer the input stage gain of the classical differential pair, and thus more emphasis may be placed on the main internal gain stage of the operational amplifier. However, for unity gain configurations, such internal gain stages are capable of providing adequate open loop gain.

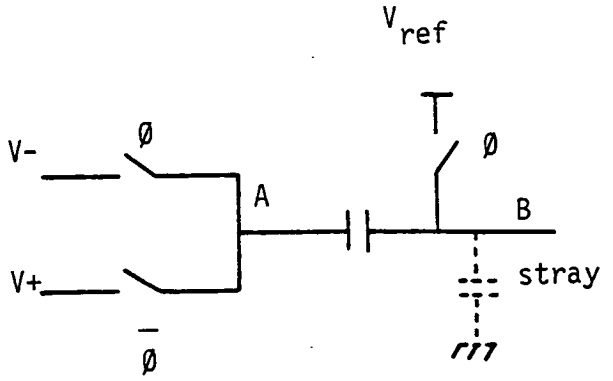


Figure 6.3: Switched-capacitor differential stage

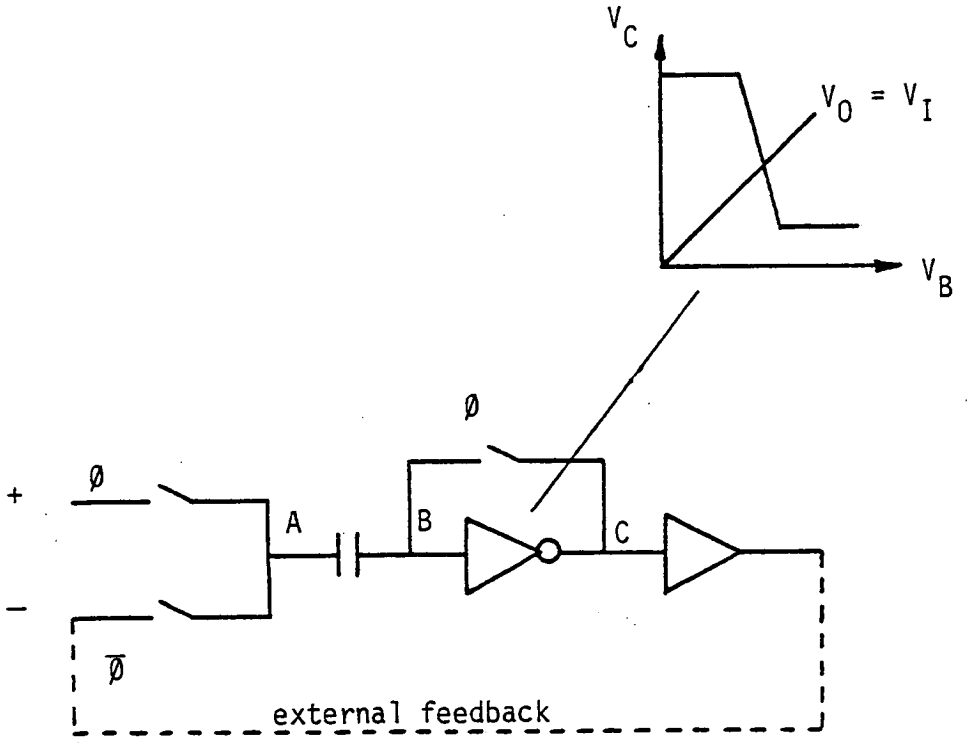


Figure 6.4: Clocked operational amplifier with internal reset

Within the complete clocked operational amplifier architecture shown in Figure 6.4 it is evident that the correct reset level for node B is that within a narrow region at the input of the invertor corresponding to the high-gain region. This is the value that node B will be forced to assume when any feedback loop is completed. Such a reset condition is achieved automatically by further exploiting the clock system, to force the condition $V_B = V_C$ around the high gain invertor during the reset period.

In a fully connected feedback configuration the amplifier operates then in two phases. On phase $\emptyset$ the + input is sampled,and the amplifier is internally reset. On phase $\overline{\emptyset}$ the feedback loop is closed and settles such that node B returns to its reset point, and therefore such that V_- is equal to the sampled value of V_+. Clearly the output is valid only during one phase and this should be chosen to suit the sampling phase of the following clocked circuitry.

By replacing the sensitive differential pair with a switched capacitor stage the design of the MOST operational amplifier is thus significantly simplified. Apart from offering this design advantage, the simplified pole pattern allows (with care) stable operation without internal compensation. These features reduce the area of the operational amplifier by a factor of 3 to typically 20,000 μm^2.

6.3 A CLOCKED OPERATIONAL AMPLIFIER FOR FEEDBACK LINEARISATION

Figure 6.5 shows the complete circuit diagram of a clocked operational amplifier suitable for the feedback linearisation of the CCD transfer function at clock rates up to 2 MHz. The basic operational amplifier block appears in the upper half of the diagram and consists of transistors M1 through M6, M11 and M12. M1, M2, M3, M4 form a single-ended cascode-connected high gain

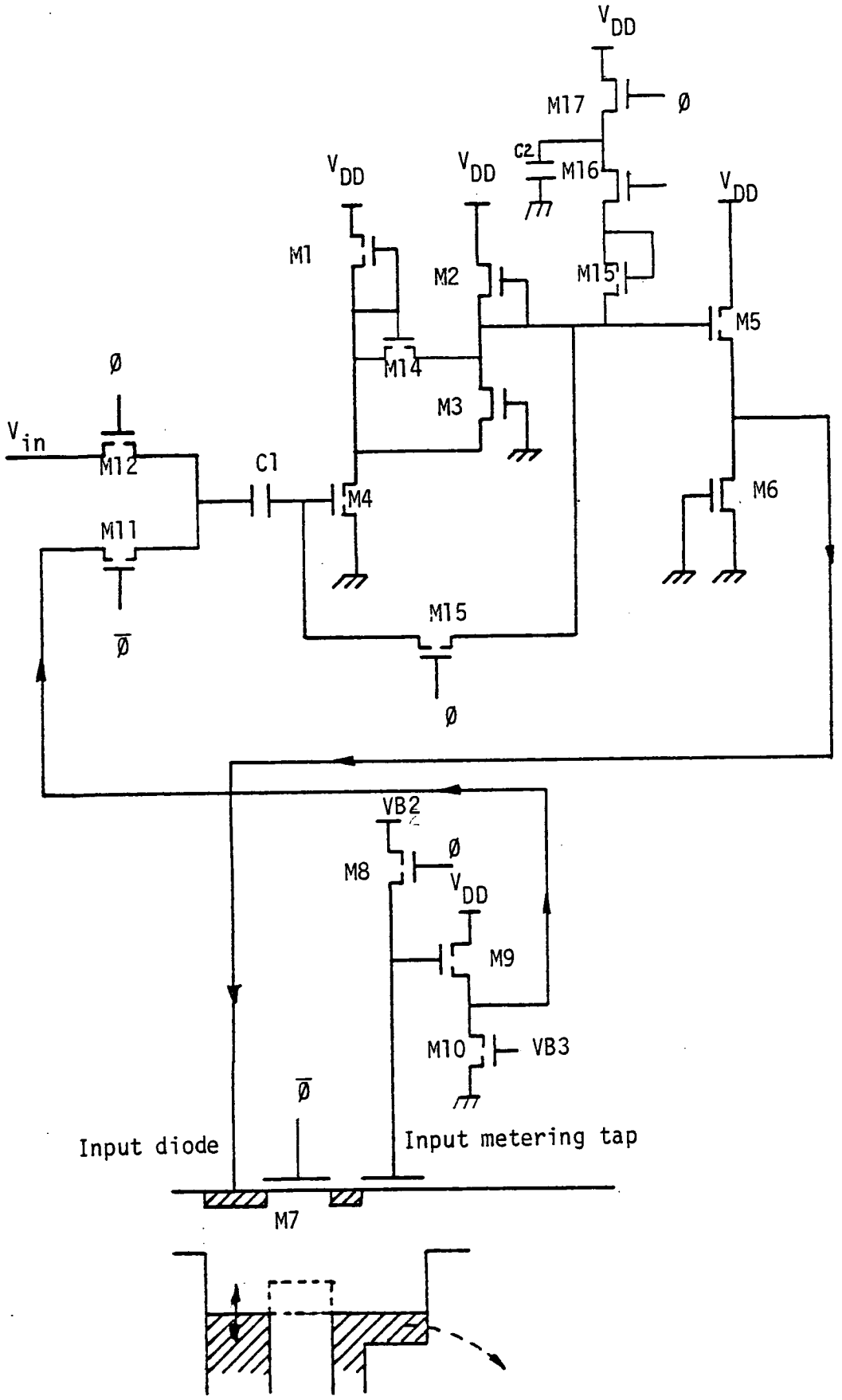


Figure 6.5: A clocked operational amplifier for feedback linearisation

Tx.	Type	L_{drawn}	W_{drawn}	Aspect Ratio
M1	P2D	10	15	2
M2	P2D	100	6	1/20
M3	P2D	10	75	10
M4	P2E	10	75	10
M5	P1E	7	90	20
M6	P2D	10	30	4
M7	P1E	7	6	1.3
M8	P1E	7	6	1.3
M9	P1E	7	45	10
M10	P1E	9	9	1.3
M11	P1E	7	6	1.3
M12	P1E	7	6	1.3
M13	P1E	7	6	1.3
M14	P1E	7	6	1.3
M15	P1E	7	6	1.3
M16	P2D	7	6	1.3
M17	P2D	7	6	1.3

Table 6.1: Operational amplifier transistor dimensions (μm)

invertor stage. M5 and M6 form an output buffer and level shifter. M11, M12 and M13 perform the necessary switching functions around the differentiating capacitor C1.

In the lower half of the diagram, M7 through M10 model the CCD input stage, with diode cut-off charge injection and charge metering via a dummy FGR tap, which is electrically identical to all subsequent taps in the CCD. This stage is non-inverting and has a linear transfer function with a gain of approximately 0.45.

Transistors M14 through M17 do not form part of the steady state loop but are included to improve the transient response of the amplifier. When the feedback loop is completed initially, the gain stage output may be too low. It can only approach the correct value through the invertor load M2, which is necessarily low-current in order to achieve a high gain. This rise time may be excessively long, and so a charge pump circuit (M15, M16, M17 and C2) is included to precharge this node to a high voltage as the feedback loop is completed. The gain stage output must then approach its final value from a higher voltage, which it is capable of doing very quickly. Indeed capacitor C2 must be chosen to limit this negative slew rate and avoid a large negative overshoot. Following this precharge operation, the diode-connected M15 isolates C2 from the gain stage output node to aid faster settling. In the event of a large negative overshoot, M14 will cut in and charge the output node some way towards its correct value through current source M1.

It should be noted that these transient-improvement functions are relevant only to peculiar nature of this feedback loop and would not normally be necessary.

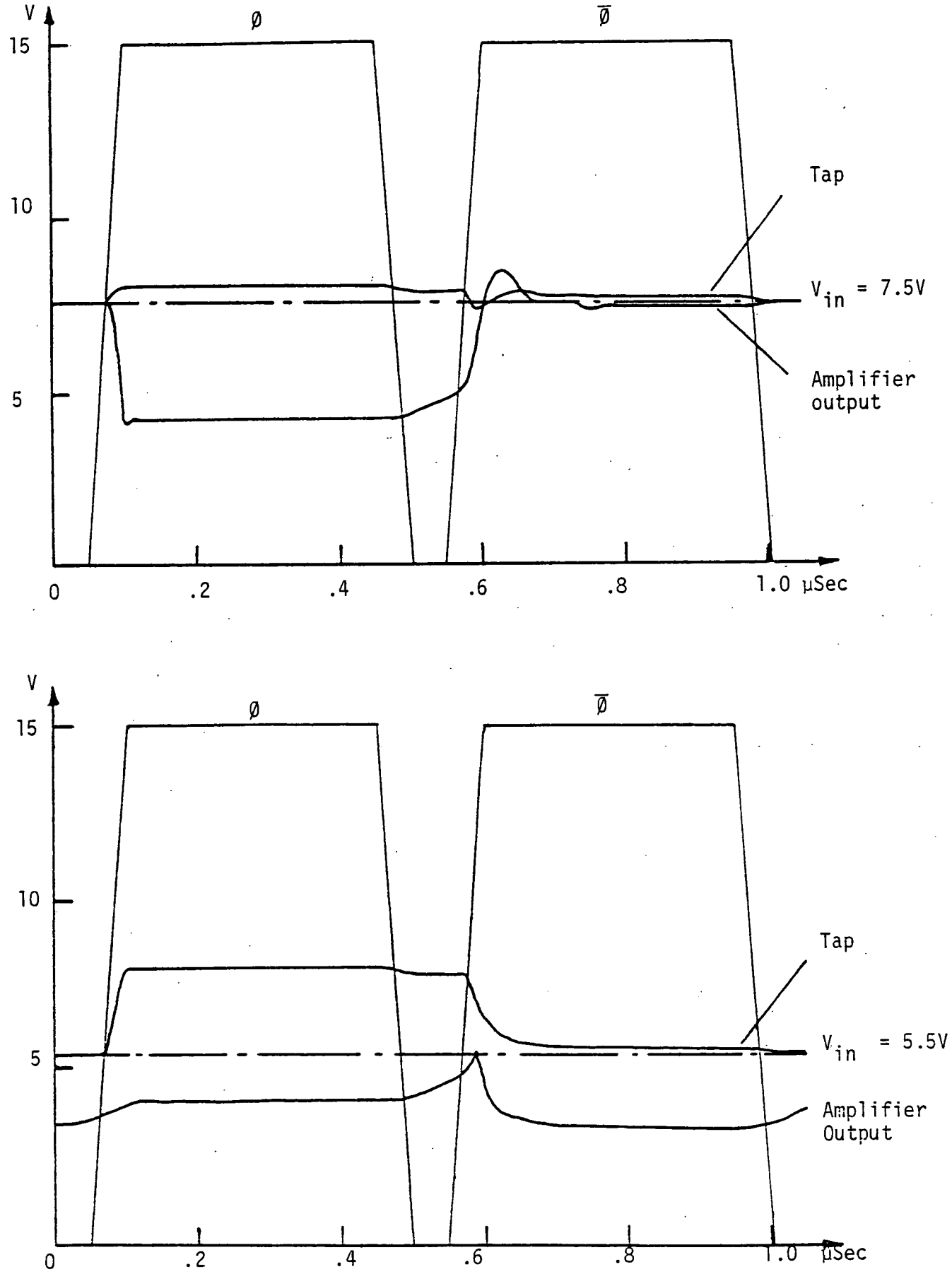


Figure 6.6: SPICE simulation of feedback linearisation

If the tap output voltage is to equal the sampled input voltage to within 1%,then a small signal open loop gain greater than 100 is required. Allowing for the various attenuating factors in the loop a gain of 300 or more is needed in the invertor stage. This is achieved using the transistor dimensions given in Table 6.1.

Figure 6.6 shows two SPICE simulations of the operation of the circuit for sampled input voltages which correspond to the maximum and minimum required tap output voltages. In this case a tap output range of 2V is required. The complementary clock phases shown represent sampling rates of 1 MHz,although the settling times will clearly allow operation at twice this frequency. Power dissipation is nominally 5 mW.

6.4 REALISATION AND OPERATION

Figure 6.7 shows a photomicrograph of a monolithic realisation of the clocked operational amplifier and feedback linearisation scheme. Note the small layout area occupied by the operational amplifier.

The two photographs shown in Figure 6.8 demonstrate correct operation of this circuit. The upper photograph shows the open-loop response of the operational amplifier block. An input ramp of 100 mV. ms^{-1} appears at the output with a slope of 20V.ms^{-1}over a range of 9.5 V using a 15 V supply and clocks. This open loop gain of 200 implies a gain of 300 within the internal invertor stage, as expected. The lower photograph demonstrates correct operation of the complete feedback linearisation scheme. The gain of the transfer function is better than .99,and linearity is better than 60 dB.

It should be noted that these characteristics are attained *automatically*, thus making the CCD a more ideal, less sensitive structure to use.

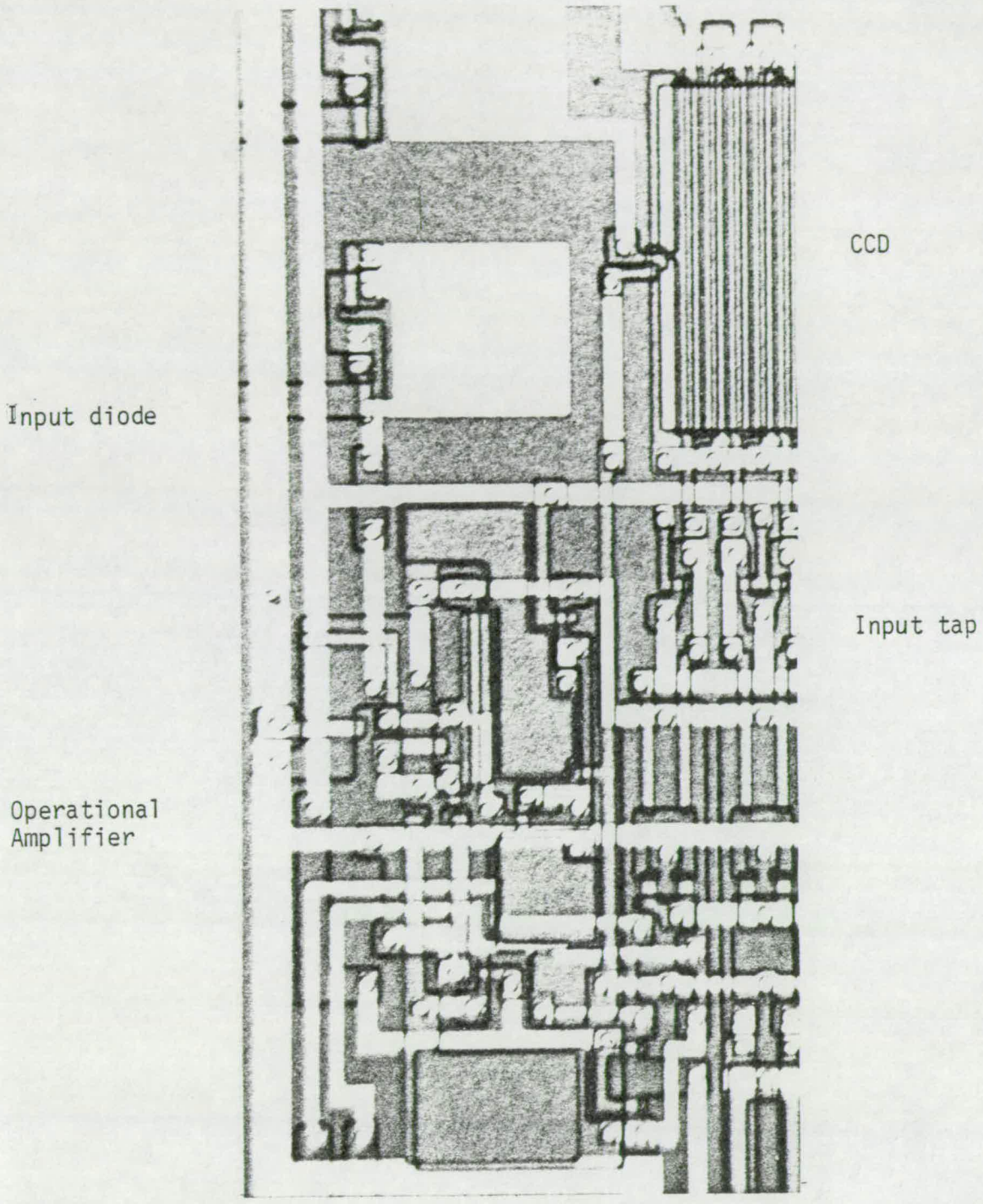
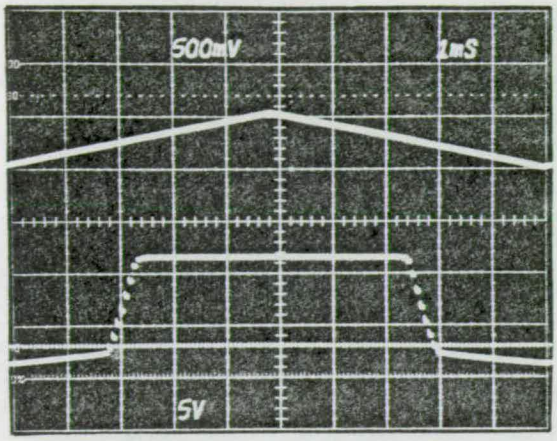


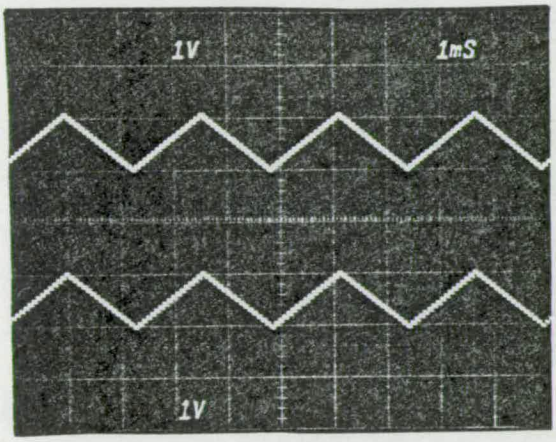
Figure 6.7: Photomicrograph of clocked operational amplifier and feedback linearisation scheme



Input

Output

Open Loop



Signal input

Metering tap output

Closed Loop

Figure 6.8: Operation of the clocked operational amplifier and feedback linearisation scheme.

CHAPTER SEVEN : A PROGRAMMABLE TRANSVERSAL FILTER (I)

The first of two practical filter realisations reported in this thesis was developed on the metal gate SET process described in Appendix 1. This process features only a single MOST threshold voltage (-1v depletion) with thin field oxides, requiring extensive 'channel-stopping'. The limitations of this technology become apparent when compared with a second filter realisation in polysilicon gate technology. Nevertheless the metal gate technology allowed the integration of 64 fully programmable filter points, in a realisation which offered the first published results from such a device[59].

7.1 CELL DESIGN

In order to make economical use of silicon area and therefore achieve a good packing density, the filter circuitry must be as simple as possible. Here the direct architecture proposed in Chapter 2, and the single transistor multiplier discussed in Chapter 4 combine to great advantage. A block diagram of the filter cell circuitry is shown in Fig.7.1. The simplicity of the structure is as important to the successful operation of the filter as to the layout packing density.

The CCD tapped delay line section is as described in section 4.3. Included within each cell are two 3-phase CCD sections, with a floating gate tap at the last electrode site. The tap is buffered to drive the gate of the multiplying transistor.

True 4-quadrant multiplication using a single transistor requires the use of a gate multiplexing scheme, and two such schemes are possible

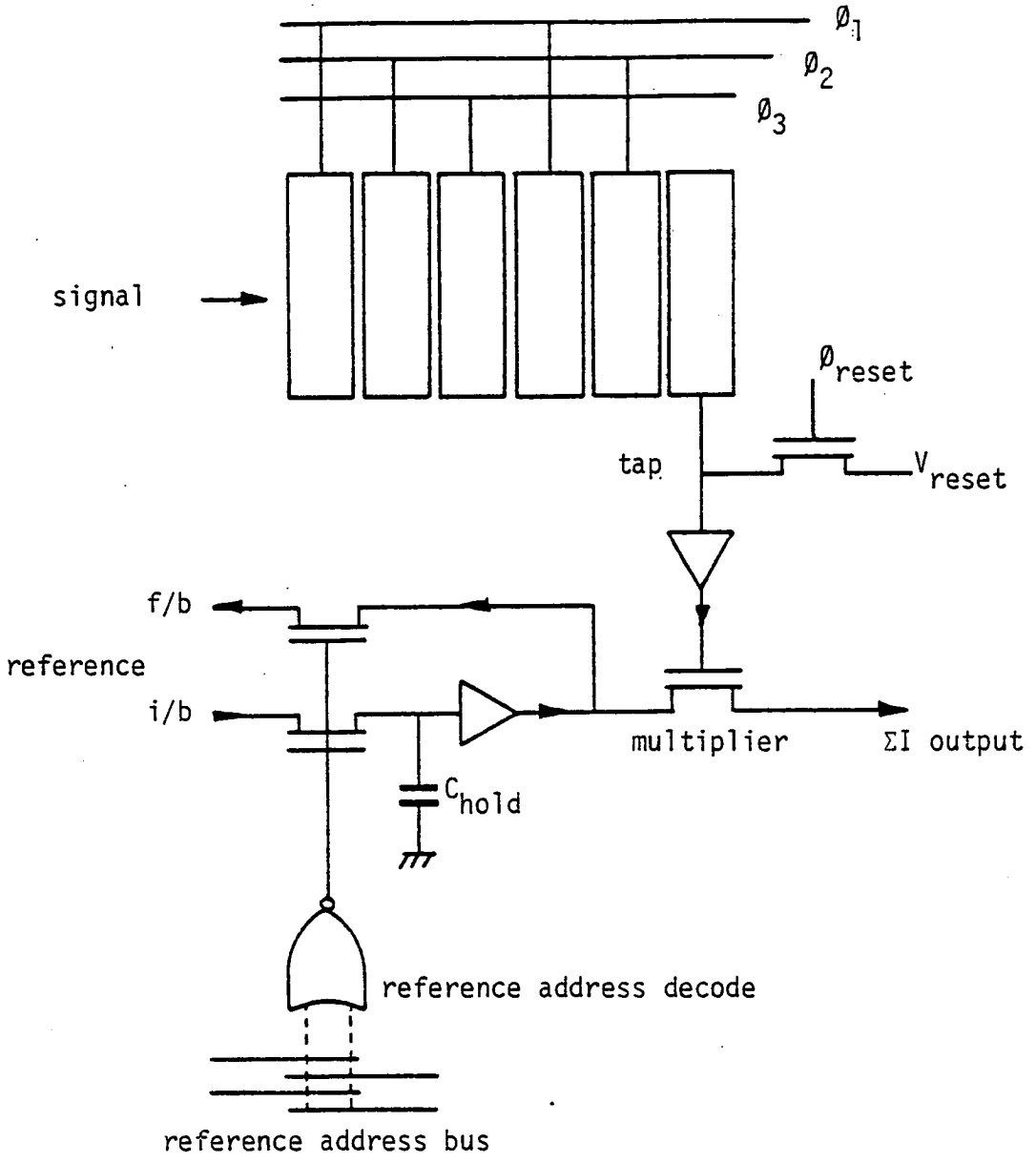


Figure 7.1: PTF I cell block diagram.

with this structure. Firstly, alternate zeroes may be propagated along the alternately tapped CCD. Secondly, the reset transistor may be used as the multiplexing element, with modulation of V_{reset} to simulate a zero sample.

On the reference side of the multiplier, a single transistor samples the analogue tap weight from the Reference Input Bus, onto a holding capacitor. This tap weight is then buffered before application to the multiplying MOST drain. A feedback loop is provided at this point through a parallel sampling transistor to sense the exact voltage applied to the multiplier after the buffer, with the intention of realising a linearisation scheme.

Finally a digital selector is provided to control the tap weight sampling gates. In this filter it takes the form of a single NOR gate which decodes a six bit address from 6 address, and 6 complemented address lines. 64 of these NOR gates are connected to form a 1-of-64 address decoder, so that any reference point may be selected for update. An essential addition to any such decoder is an *enable* line gating all outputs to avoid falsely decoding any spurious address state, which may occur whilst the address is changing.

A detailed transistor circuit diagram is given in Figures 7.2 and 7.3 with corresponding transistor dimensions in Table 7.1. This filter, being regarded as an early prototype, allows a maximum of flexibility. Thus all supply and bias voltages are brought out separately for individual control. This feature makes the device apparently complex to drive. However the second filter design, to be described, is fully committed to a single supply with all biasing generated on chip, removing this adverse complexity.

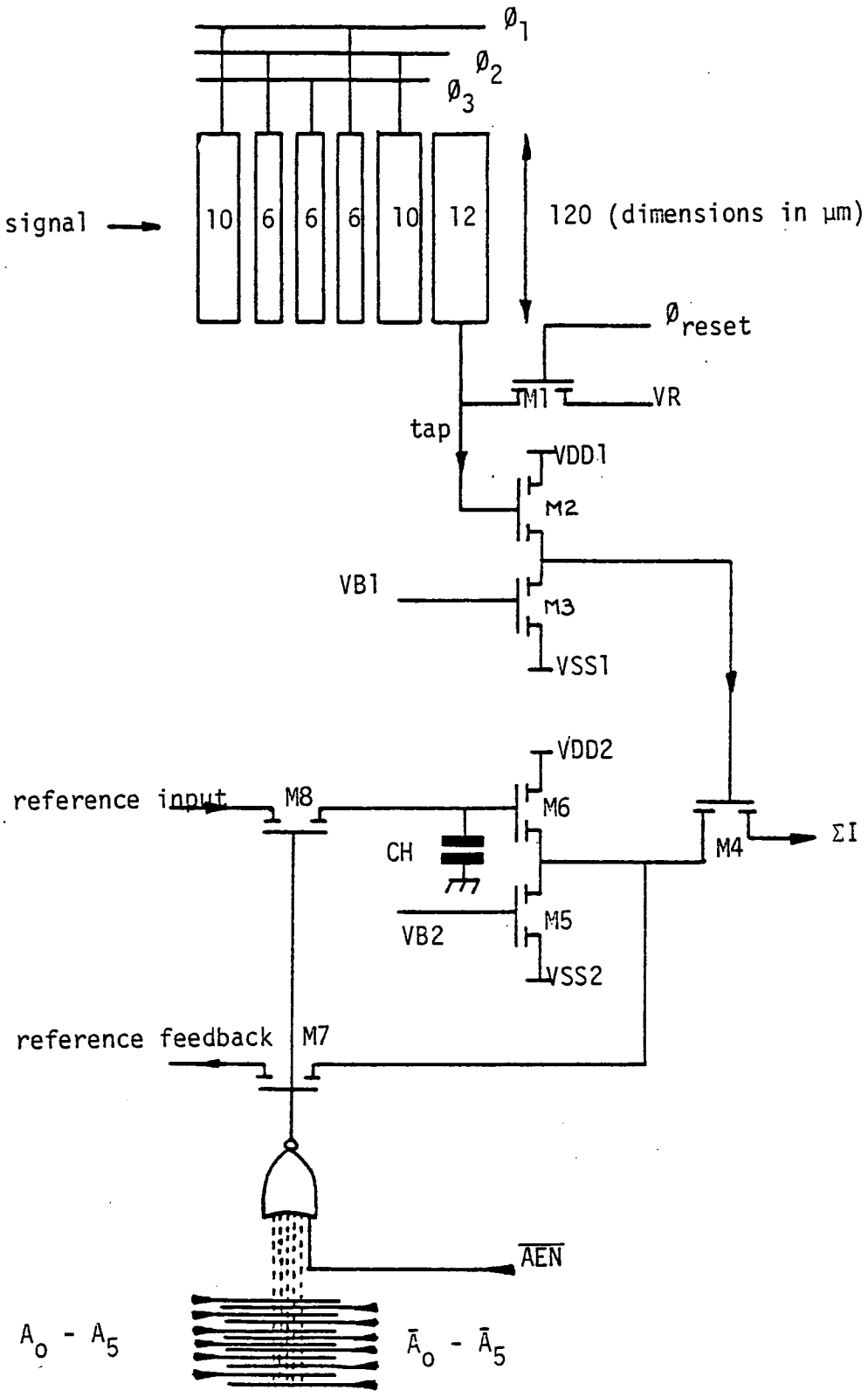


Figure 7.2: PTF I cell circuit diagram

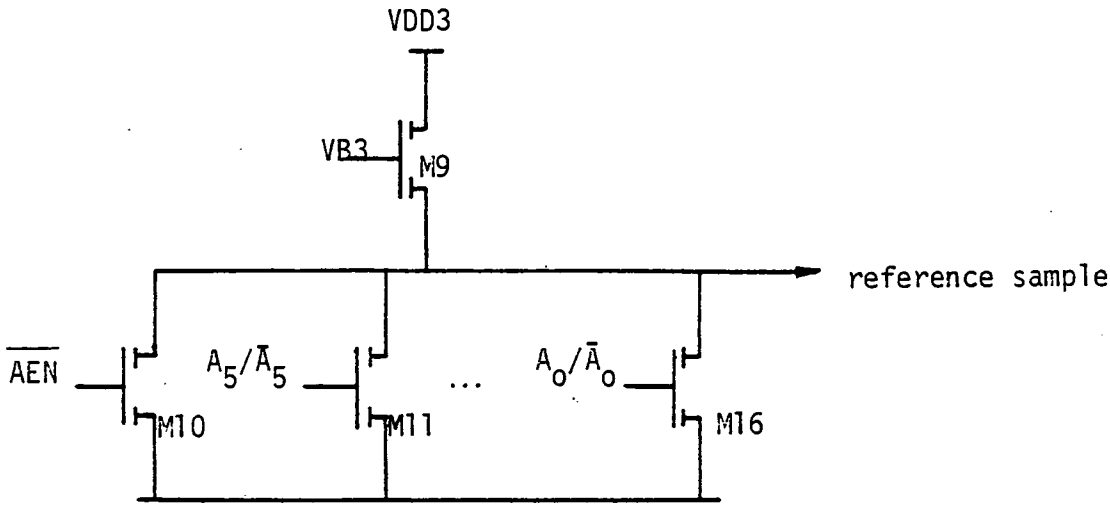


Figure 7.3: PTF I reference address decode gate

TX	DRAWN W/L	ACTUAL* W/L	ESTIMATED β μA/V
M1	9/10	7/7	15
M2	72/10	70/7	150
M3	9/10	7/7	15
M4	12/153	10/150	1
M5	9/10	7/7	15
M6	72/10	70/7	150
M7	72/10	70/7	150
M8	9/10	7/7	15
M9	5/93	3/90	0.5
M10-16	22/23	20/20	15

* Allowing 1 μm edge reduction metal I.

1.5 μm sideways diffusion N^+.

Table 7.1: PTFI CELL TRANSISTOR DATA

7.2 DELAY AND ADDRESSING STRUCTURE

Apart from the repeated filter cell, the complete device must also allow for some CCD input stage, and for access to a final CCD tap for cascading. For simplicity these are the only additional features provided on this design.

A diode cut-off CCD input stage with floating gate metering tap is provided, as described in section 4.2. The output from this tap is made available for monitoring purposes, and to allow for feedback linearisation using an external operational amplifier.

A serial signal output for cascading is provided, simply by monitoring the output of tap 64. After this tap, the charge in the CCD is transported to a destructive output diode.

The resultant delay and address structure for the cascadable 64 point filter is shown in Fig.7.4.

7.3. CHIP LAYOUT AND ASSEMBLY

Since the CCD must be tapped on a regular short pitch, the filter cell should ideally be realised as a thin cascadable strip. The layout geometry of this cell is dominated on the metal gate process by the necessary separation of transistors with guard rings of p+ channel-stop. This feature constrains the minimum cell width to 56μm, whilst the cell depth emerges as 2056 μm. The layout for this cell is given in Appendix III.

The 64 point filter is thus formed as a single block of cascaded cells. A photomicrograph of the completed chip, measuring 4.5 x 3.3mm (179 x 130 mils), is shown in Fig.7.5. Visible in the upper portion of the device is the CCD tapped delay line, and along the bottom,

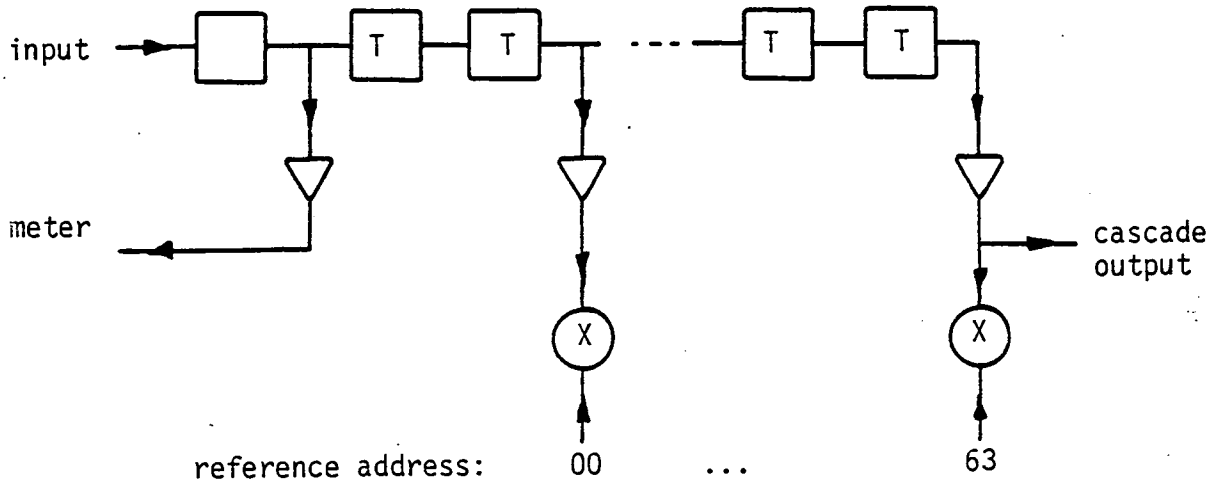


Figure 7.4: PTF I delay and address structure

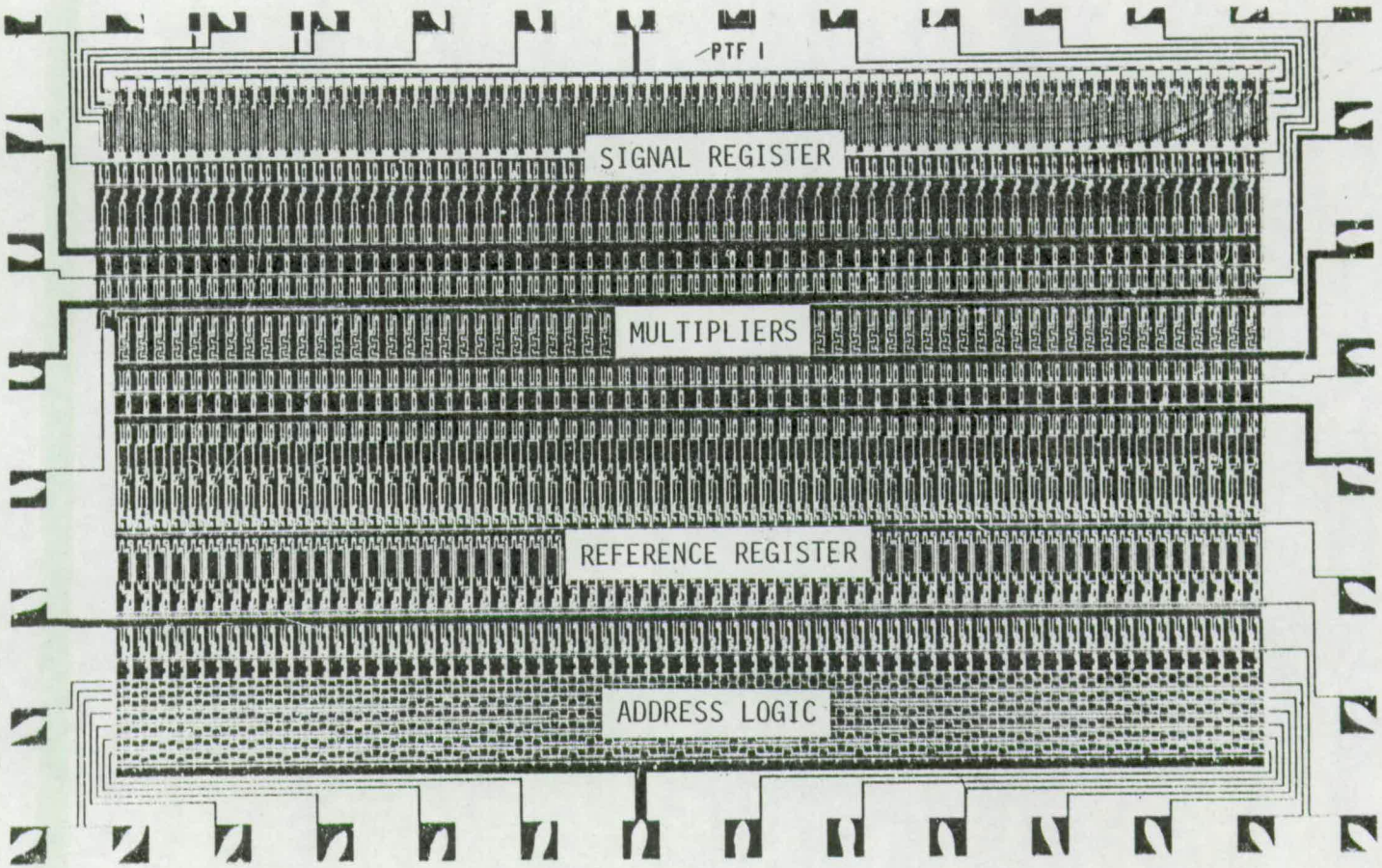


Figure 7.5: Photomicrograph of PTF I Device

the reference address decoding logic. All connections to the filter block are bonded out directly, in keeping with the flexible design philosophy. The chip is assembled for use in a standard 40 pin package.

7.4 DEVICE CHARACTERISATION CIRCUITS

Some peripheral circuitry is naturally required in order to test the filter. The flexibility of this design however results in a somewhat complex peripheral requirement, in that many of the functions that might normally be integrated with the filter must be attended to externally. These functions include the generation of multiple clock waveforms (to drive the CCD), bias and supply voltages, and address signals, all level shifted to suit the internal circuitry.

A prototype filter driving circuit and timing diagram is given in Appendix IV. Although the operation of the circuit is not considered here in detail, several features are worthy of note. The CCD is driven from a buffered 3-state counter, with clock waveforms between +2V and +25V referred to the chip substrate. The floating gate taps are reset to 15V. V_{SS} supplies are nominally all at +2V to allow correct field-plate action around the transistors. V_{DD} supplies to the linear circuitry are at +15V, and to the digital circuitry at 20V. Address signals are supplied to the decoder from a CMOS counter between 0V and 15V. The signal input range is 8-10V, with tap weights in the range 4-8V referred to a virtual earth of 6V at the filter output.

An operational output summing amplifier scheme with current feedback, as shown in Fig.5.2 (b), has been implemented.

The results to be shown here are those obtained from optimised driving circuitry, which is the result of several design iterations.

Two improvements to performance obtained during this process are of particular note. Firstly, because of the sensitivity of the multipliers to common changes in the applied reference voltages, it is important that as few timing events as possible occur between the '*zero*' and '*signal*' phases of operation. Secondly, it is preferrable that the signal zero level be separately switched onto the multiplier gates through the reset transistors, rather than via alternate zeroes transmitted along the CCD. This is because alternate zeroes in the CCD become corrupted by charge transfer inefficiency and so they become less effective at later CCD taps. Forcing the zeroes through the reset transistors ensures a good level at all multiplier gates and further, allows the CCD to operate in double sampling mode to improve the effective charge transfer efficiency.

The reference linearisation scheme has not been implemented, since the tap weight accuracy appears to be satisfactory without this facility.

The driving circuit transforms the device into a true black-box programmable transversal filter. To characterise this filter the experimental arrangement shown in Fig.7.6 has been adopted. Two purpose-built signal generators supply the signal and reference channels, with the filter output monitored by oscilloscope or, where appropriate, by spectrum analyser. The signal channel source is a digital function generator which may be programmed to supply repeated sinusoidal and square waves, as chirps or discrete frequencies. The reference channel source repeatedly generates a tap weight vector, to programme and refresh the on-chip reference. This tap weight vector can be loaded from the signal source for autocorrelation or matched filtering, or manually programmed to achieve any desired impulse response.

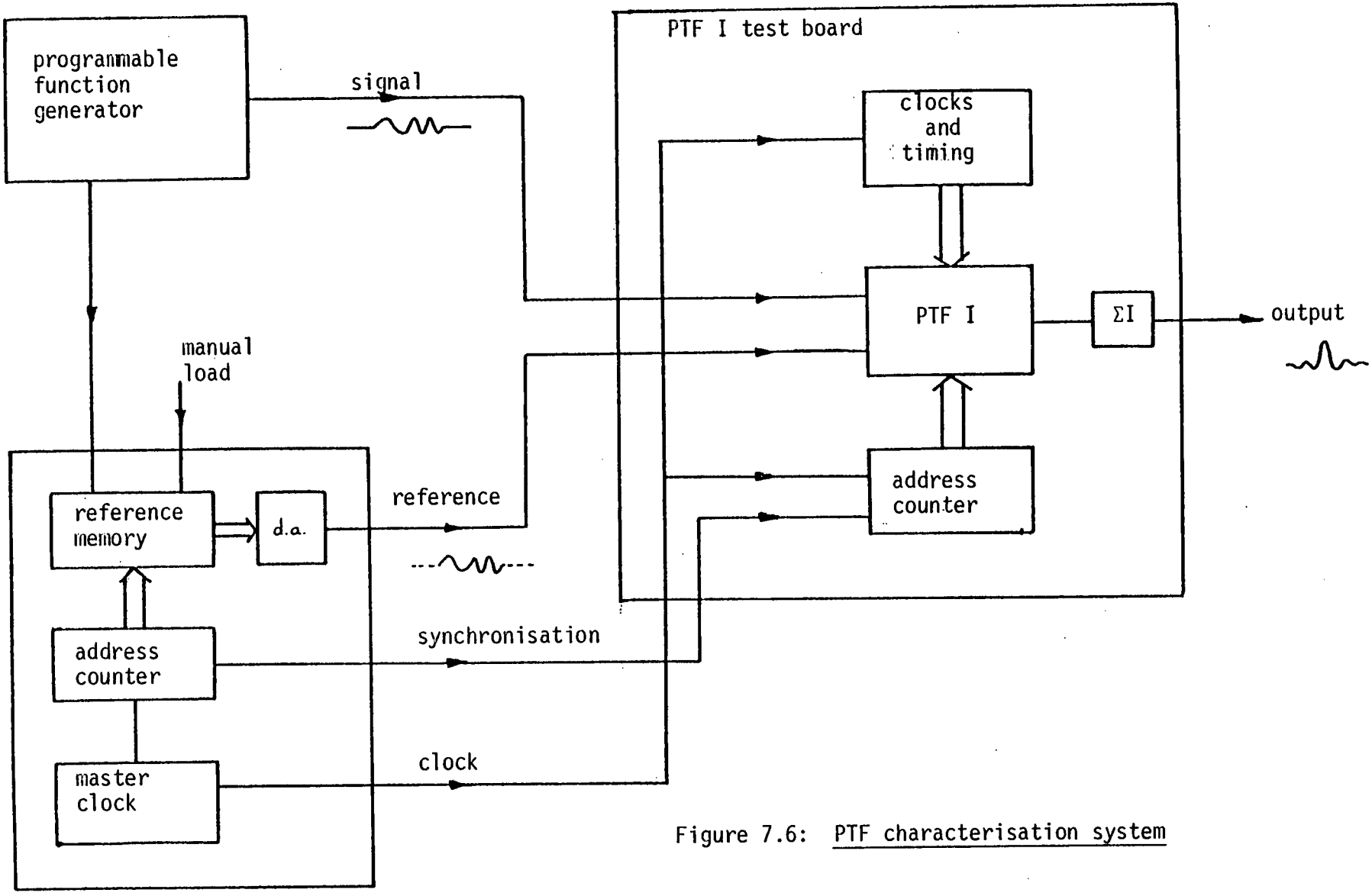


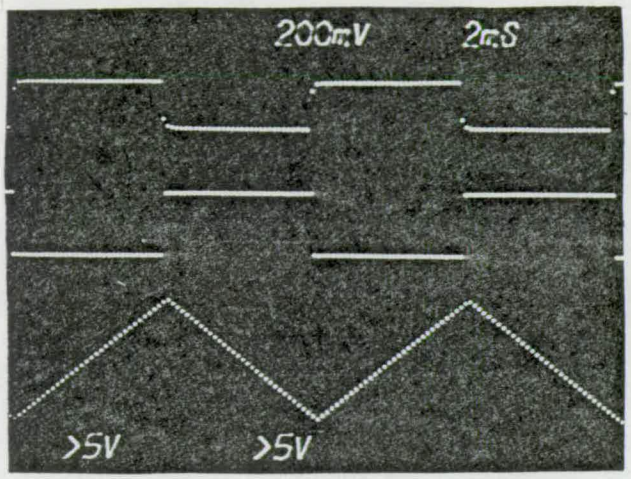
Figure 7.6: PTF characterisation system

7.5 RESULTS

An elementary test of the programmable filter is a correlation of two squarewaves, which should ideally yield a triangular waveform. Fig.7.7 demonstrates the autocorrelation of a squarewave of period equal to the length of the filter. The upper trace shows the delayed CCD signal output at tap 64, demonstrating the residual effects of charge transfer ineffieciency. The centre trace is the applied tap weight reference waveform, and the lower trace demonstrates clearly the expected triangular output. At 20V p-p this output uses the full available dynamic range of the circuit, since all filter points are fully saturated. The random output noise component is estimated at 6mV r.m.s., giving a peak output dynamic range of 70 dB. Note the quality of the CCD output waveform, which is available for cascading, and the linearity of the output triangle.

A more stringent test of device accuracy and dynamic range is to examine the impulse response, a typical example of which is shown in Fig.7.8. Here the programmed reference coefficients represent a Hamming-weighted sinc x function, used to realise a low pass filter response. The impulse test is a good indication of performance because only one filter point is active at any time, while all filter points continue to contribute noise. The measured dynamic range of this impulse response is 34 dB.

The frequency response of this low-pass filter realisation is shown in Fig.7.9 (a). Evidently a successful characteristic has been obtained. From this figure the average stopband attenuation is estimated at 34 dB, with a minimum value of 26 dB. Since the theoretical stopband level is approximately -50 dB, some error is apparent. This error may be attributed to tap weight innaccuracy[60] and distortion,



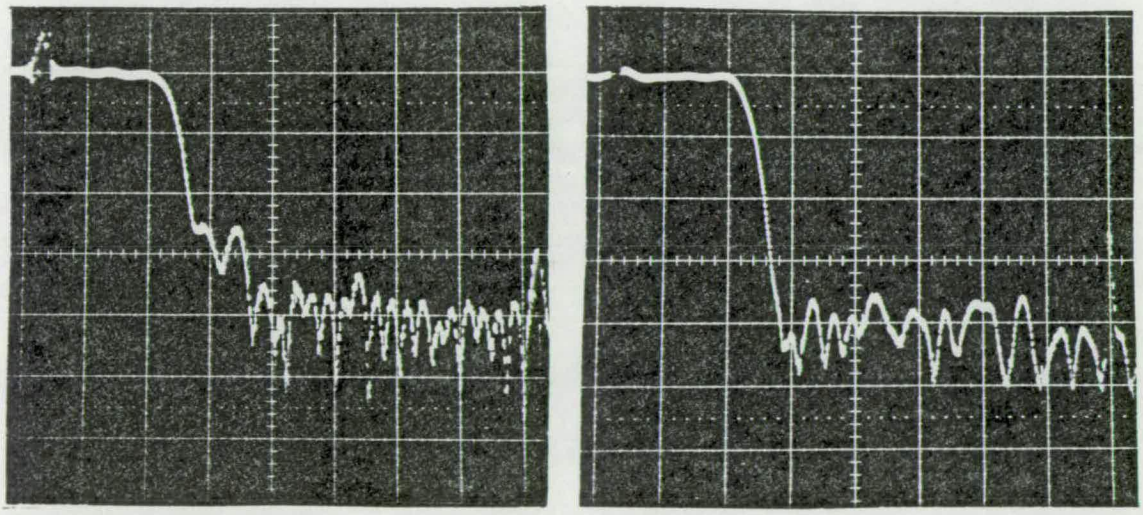
tap 64

reference

filter output

Figure 7.7: Square wave correlation f_c = 6kHz

Figure 7.8: Hamming weighted low pass filter impulse response



(a) (b)

Figure 7.9: Frequency responses obtained from 7.8.

(a) Tap weights directly applied
(b) After manual correction

Horizontal: 1 kHz/div
Vertical: 10dB/div

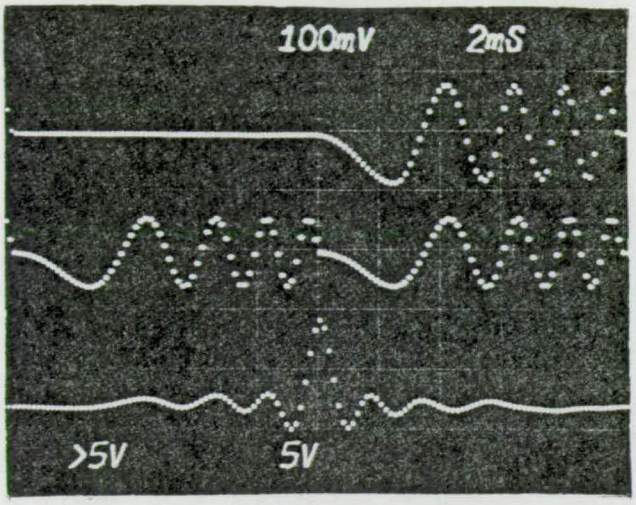
which are considered further in Chapter 9.

Now in a stationary-weighted filter it is possible to correct for tap weight errors, to arbitrarily improve the filter response. Such an improvement is demonstrated in Fig.7.9 (b), which exhibits an average stopband attenuation of 40 dB (minimum value 34 dB).This correction, applied manually here, is indicative of the performance that may be achieved when some degree of feedback is allowed in setting the tap weight coefficients.

Of course the prime application of this device is as a matched-filter, or correlation detector. Fig.7.10 shows the matched filtering of a chirp waveform sweeping from d.c. to f_c/8Hz. The correlation peak is of the expected sinc x form, demonstrating a compression of the signal energy. If the sweep is increased to the Nyquist limit then the output is compressed into a peak only one sample wide. This is demonstrated in Fig.7.11. Again the quality and symmetry of the output waveform are of note. The effects of charge transfer inefficiency are evident in the CCD output, where the higher frequencies are slightly attenuated, and in the filter output where an assymmetry is barely perceptible at the base of the peak.

Many other functions have been programmed and tested, including for example, a differentiator (Fig.7.12), and a bandpass filter (Fig. 7.13). Further demonstrations of the versatility of the filter are limited only by the imagination of the user.

Table 7.2 summarises the major performance parameters of this prototype device. Factors such as charge transfer inefficiency and dark current leakage are essentially process-determined, although the double-sampling technique certainly improves the former. Noise is the most difficult parameter to control, as it often stems from

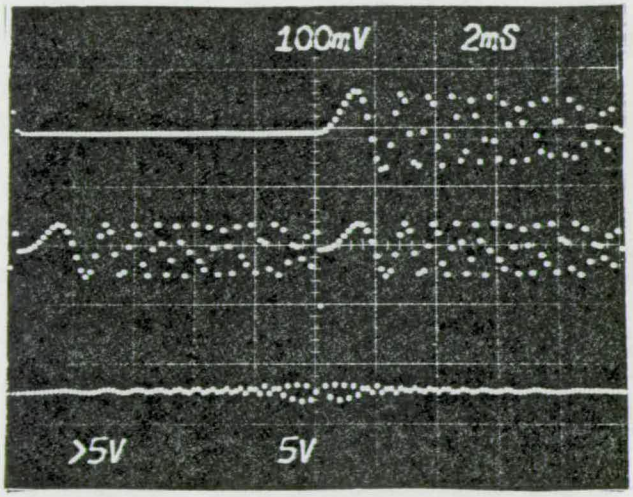


tap 64

reference

filter output

Figure 7.10: Chirp autocorrelation 0 - $f_c/8$



tap 64

reference

filter output

Figure 7.11: Chirp autocorrelation 0 - $f_c/2$

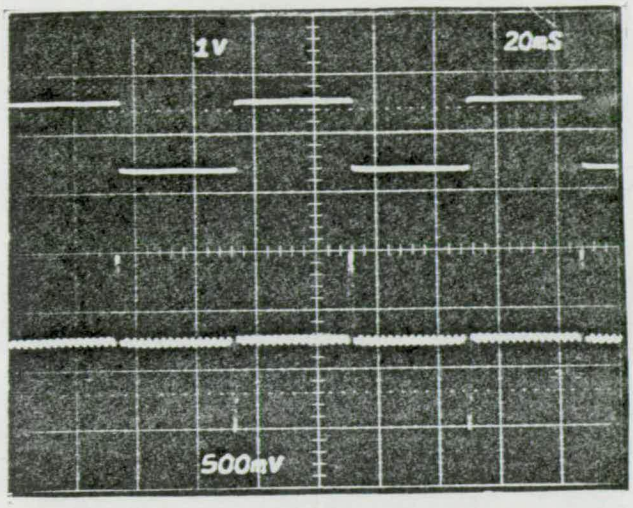


Figure 7.12: Differentiator

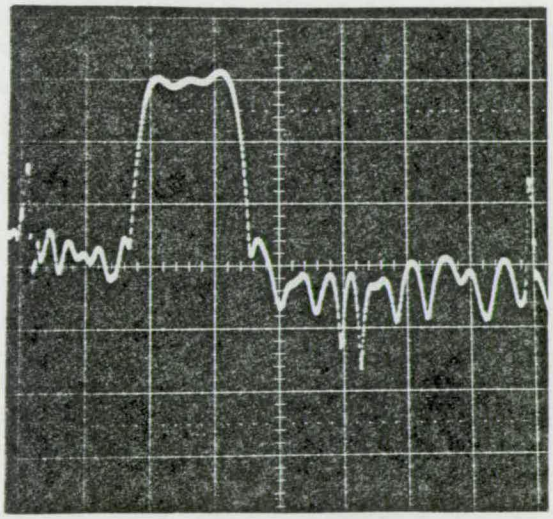


Figure 7.13: Bandpass Filter

Horizontal: 1 kHz/div
Vertical: 10dB/div

Number of Filter Points		64
Power Dissipation	(1)	5 mW/point
Cell Area		2.2 x 120 mils/point
Charge Transfer Inefficiency	(5)	10^{-3} per point
Harmonic Distortion	(2)	-34 dB
CCD Through-gain		$\approx .5$
Clock Rate	(3)	500 Hz - 100 kHz
Dynamic Range (unweighted up-chirp correlation)		64 dB
Dark Current	(4)	20 nA cm^{-2}
Reference Decay	(4)	1 $V \cdot sec^{-1}$

Notes: (1) Typical

(2) Signal 1v p-p, reference $\pm$1v

(3) System limited

(4) Room temperature

(5) Double Sampling Scheme, measured at tap 64.

Table 7.2: PTF I Performance Summary

the driving circuitry rather than the device itself; indeed most of the prototype development time was spent on this aspect. Again circuit techniques can help, and in this case the chopper action in the multiplier is extremely useful, as becomes apparent in the following chapter.

The maximum useable clock frequency of 100 KHz is limited by the complexity of the clocking system; again a device which is simpler to drive might be expected to realise greater bandwidths.

Despite the inevitable limitations of a prototype design, this device appears well suited to the sonar matched filtering application.

7.6 APPLICATIONS

Following the success of the 64-point programmable filter many applications have evolved. A major application is in the sonar system for which it was developed. Four of the chips are cascaded to form a 256-point programmable matched filter on a double-size Eurocard, complete with refined driving circuitry, as shown in Fig.7.14. This card has undergone successful tests on real sonar returns and compares favourable with an 8-bit digital benchmark.

The 64-point device has also formed the basis of some experimental work on Adaptive Filters[96], which automatically adjust the tap weight coefficients by feedback to achieve desired filter responses. The random reference address capability, hitherto undemonstrated, has been successfully exploited by Kapur[97] in a novel spectrum analyser. Both of these developments have spurred interest in dedicated architectures based on this device.

The single negative factor preventing further exploitation of this device is the obsolescance of the metal gate process. The limited

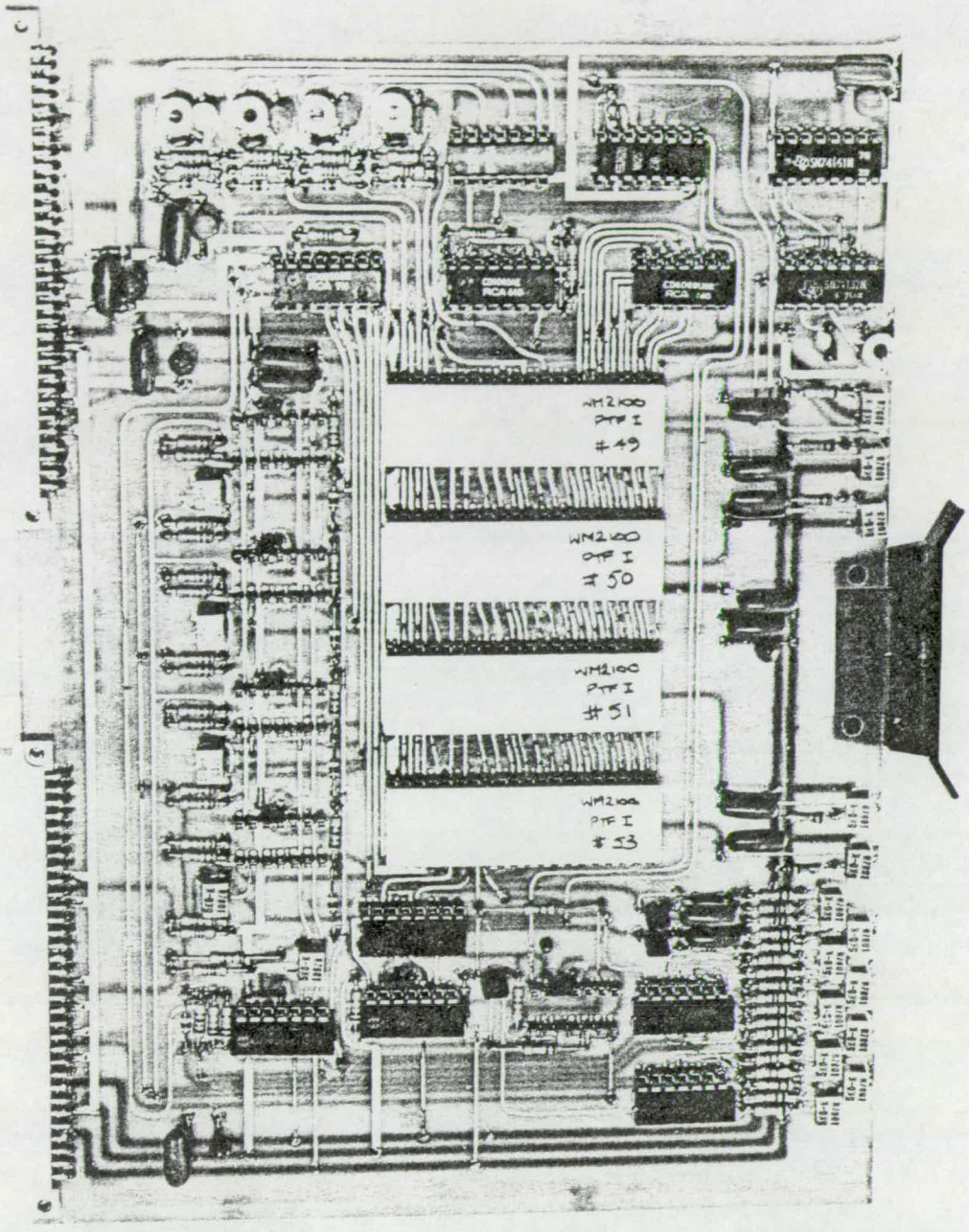


Figure 7.14: 256-point matched filter module

(Courtesy Wolfson Microelectronics Institute)

availability and lifetime of the prototype devices has restricted their further development, but this should not detract from their significance. The performance parameters still exceed in many aspects those of other reported developments.

CHAPTER EIGHT : A PROGRAMMABLE TRANSVERSAL FILTER (II)

A second filter realisation, PTF II, was developed during 1978/9 on the double polysilicon gate process also described in Appendix I. The filter is developed upon the principles successfully demonstrated on the PTF I design. In this case however, every effort has been made to simplify the use and operation of the device. This philosophy is partly born out of the process characteristics, which offer 15V clocking, and a range of threshold voltages to simplify interfacing, and partly out of design, which includes on-chip feedback linearisation to improve the CCD transfer function, and shift register controlled reference update to reduce pin count. The multiplier has been reduced to non-multiplexed 2-quadrant operation, which is sufficient for matched filtering against a stationary reference. The device is assembled for use in a standard 16-pin DIL package.

An immediate consequence of the adoption of a self-aligned thick film process is a significant reduction in the dimension of the filter cell layout. The disappearance of channel-stop problems greatly reduces the average silicon area occupied by each transistor, and the self-aligned gate property allows the use of mirror symmetry techniques within the cell layout without the need for attention to overlap tolerances. These two features combine to reduce the filter cell pitch to 28μm, exactly half that achieved on the metal gate process. A linear reduction factor of 2 implies a reduction in area by a factor of 4, thus a 256 point filter becomes feasible within the area previously occupied by the 64 point design.

8.1 CCD DESIGN

A filter cell pitch of 28μm leaves no choice in CCD architecture. Indeed this corresponds to the minimum achievable pitch for a single 2-phase CCD stage. The CCD structure then is exactly as described in Section 4.4. The four electrode cell comprises two split-phase clocked electrodes, a single biased electrode and an FGR tap.

8.2 CELL DESIGN

The linear MOS cell architecture shown in Fig.8.1 is based strongly upon the circuitry developed for PTF I. Again a tap buffer drives the gate of the MOST multiplier, and the tap weight is sampled from a Reference Input Bus onto a holding capacitor and applied to multiplier MOST drain via a buffer. Two changes are evident. Firstly a sample-hold transistor is included at the multiplier gate. Since multiplexed operation is not intended, this is a simple method of ensuring a continuous sampled and held filter output, free from spurious transisions. The second circuit simplification is the removal of the reference feedback facility, reliance is now placed upon accurate buffer matching along the tap weight register. However it is anticipated that this matching tolerance will be more than adequate for the major matched filter application intended.

Evident from Fig.8.1 is an important difference in the method of tap update selection. During extensive use of PTF 1 it was found that tap updates were invariable performed sequentially, and this is certainly a natural consequence of using an external tap weight RAM for refresh. Thus the externally cumbersome address decode system is here replaced by a single dynamic shift register, requiring only a serial input and output,and a two-phase complementary clock. In principle

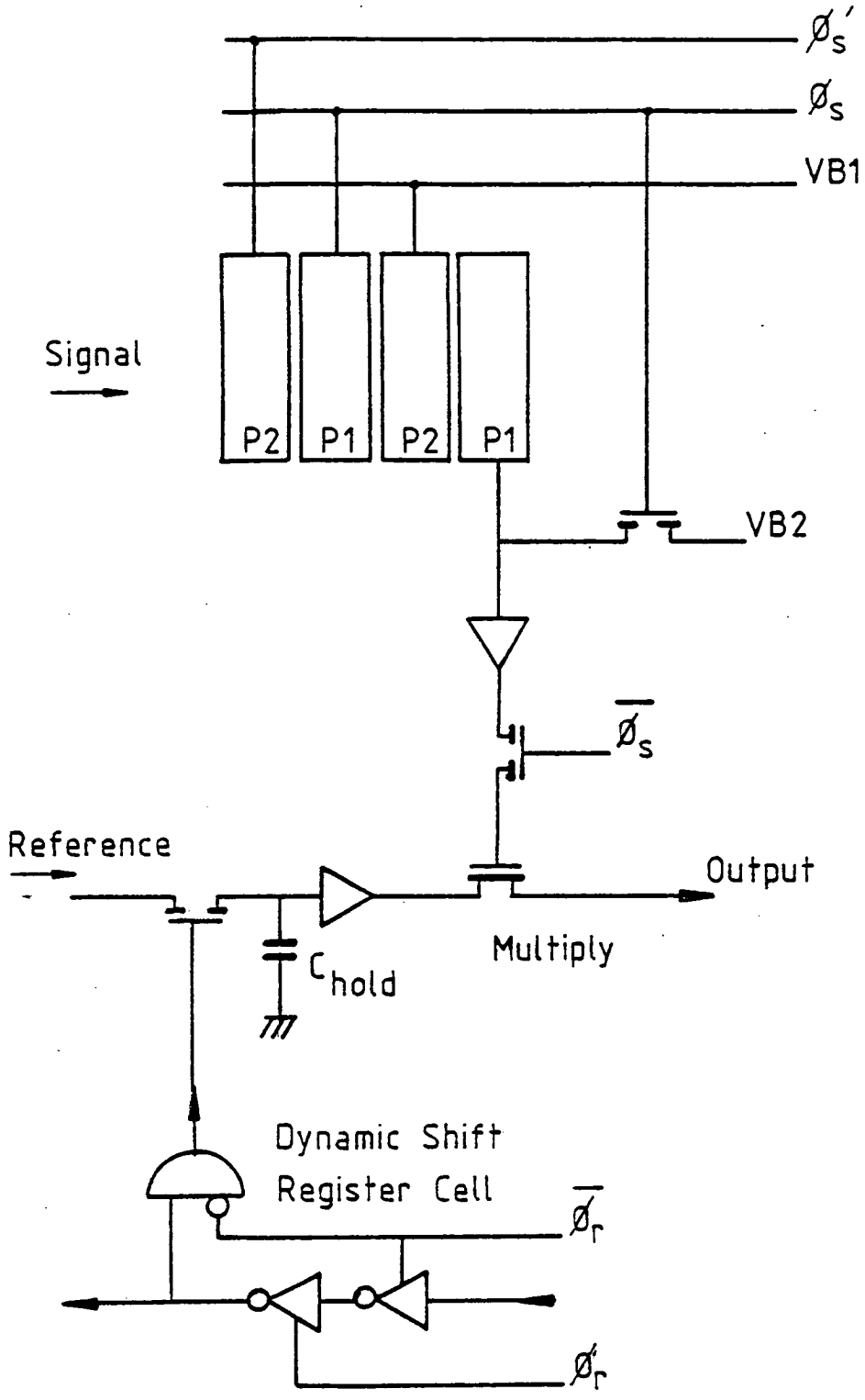


Figure 8.1: PTF II cell block diagram

a digital '1' is entered at the serial input and this is clocked along the register, opening the sampling transistor at each tap weight store in sequence. The '1' finally appears at the serial output for cascading. The direction of shift is deliberately chosen to be opposite to that of the CCD signal register, thus realising a *correlation* of the entered tap weight vector with the sampled signal vector.

Gating is provided at the parallel outputs of this Reference Shift Register to ensure that successive 1's do not overlap at adjacent cells. This is a similar problem to the previous one of decoding spurious address states and is overcome without the need for a separate enabling line, by simply gating the outputs with one of the clock phases.

Because the tap weights must be continually refreshed the shift register may be dynamic, and this leads to a significant simplification in design.

The dynamic cell used is standard, save that the inverter stages themselves are powered from the clock lines. This not only reduces power dissipation by virtue of the reduced duty cycle, but is a necessary measure if these stages are to be incorporated within the 28μm cell pitch, since there is not room for a separate supply rail.

Timing requirements for the new filter cell are altogether less critical, as may be seen from Fig.8.2. Apart from the split phase necessary for the CCD, one other (complementary) phase is required for the tap signal sampling transistors. The Reference Shift Register requires two simple, non-overlapping complementary clock phases which may be separate from the signal timing or may, for convenience, be commoned to two of the signal phases.

Figure 8.3 shows the full filter cell schematic, with transistor

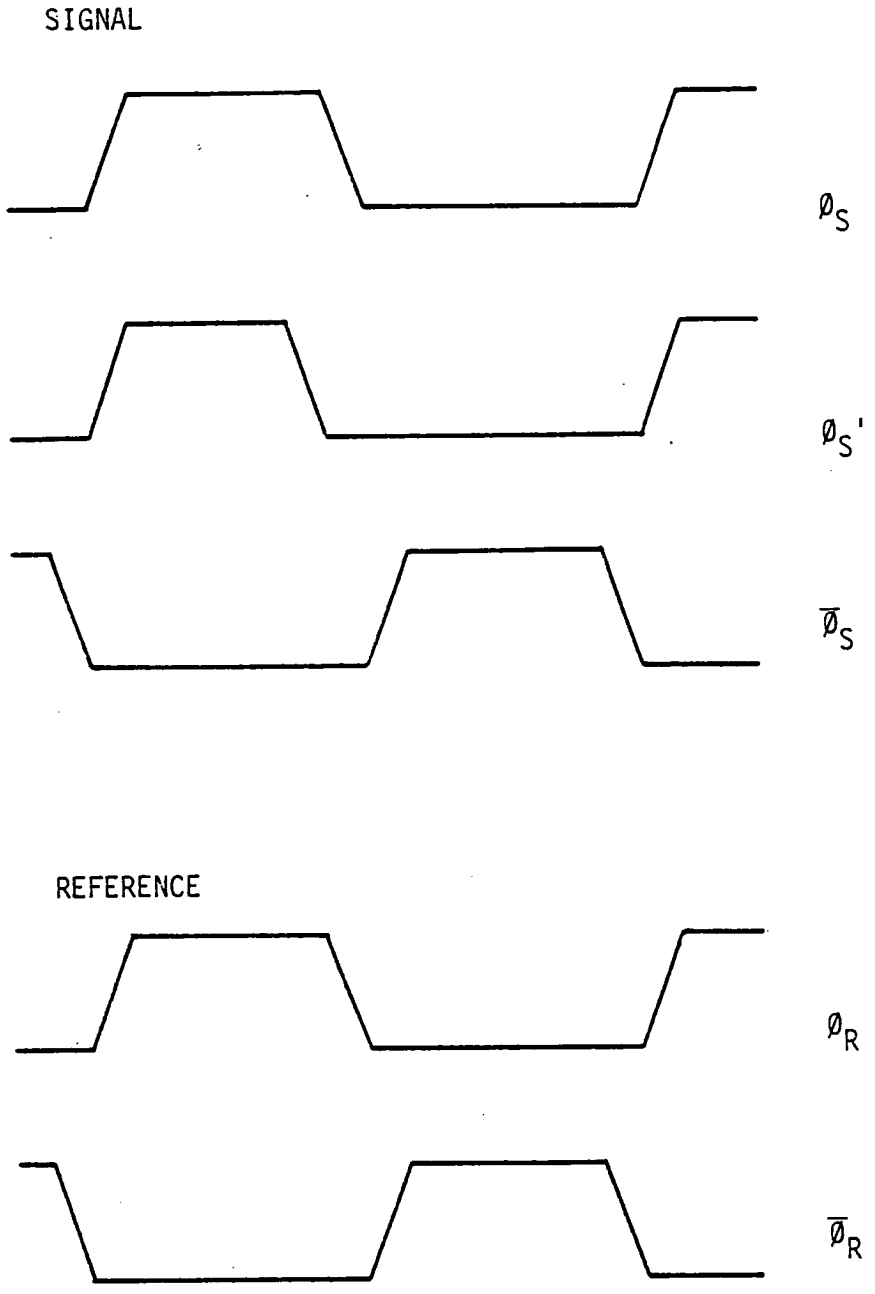


Figure 8.2: PTF II timing

dimensions given in Table 8.1. In detail;

M1, M2, M3 and M4, M5, M6 form a 2-phase dynamic shift register stage. High-threshold active transistors are chosen to maximise noise immunity.

M8 clears any 'enable' during ϕ_R in order to avoid overlapping enables on successive reference points. M7 isolates this clearing operation from the shift register output holding node.

M9 is the tap weight sampling transistor, with C_H as a 1 pF holding capacitor.

M10, M11 and M14, M15 form the tap weight and tap signal buffers respectively. The transistors used here are exclusively those which demonstrate the best gain and threshold matching (that is those transistors that undergo the minimum of implantation during processing)

M12 performs the multiplication and here a deep depletion transistor type is chosen, to ensure a maximum effective gate drive, and thus minimise signal distortion through the multiplier (section 5).

M13 samples the tap signal buffer output onto the gate of M12, which acts as a parasitic holding capacitance.

M16 is the tap reset transistor.

Note that wherever possible, supply and bias voltages are commoned throughout the cell to minimise external connections.

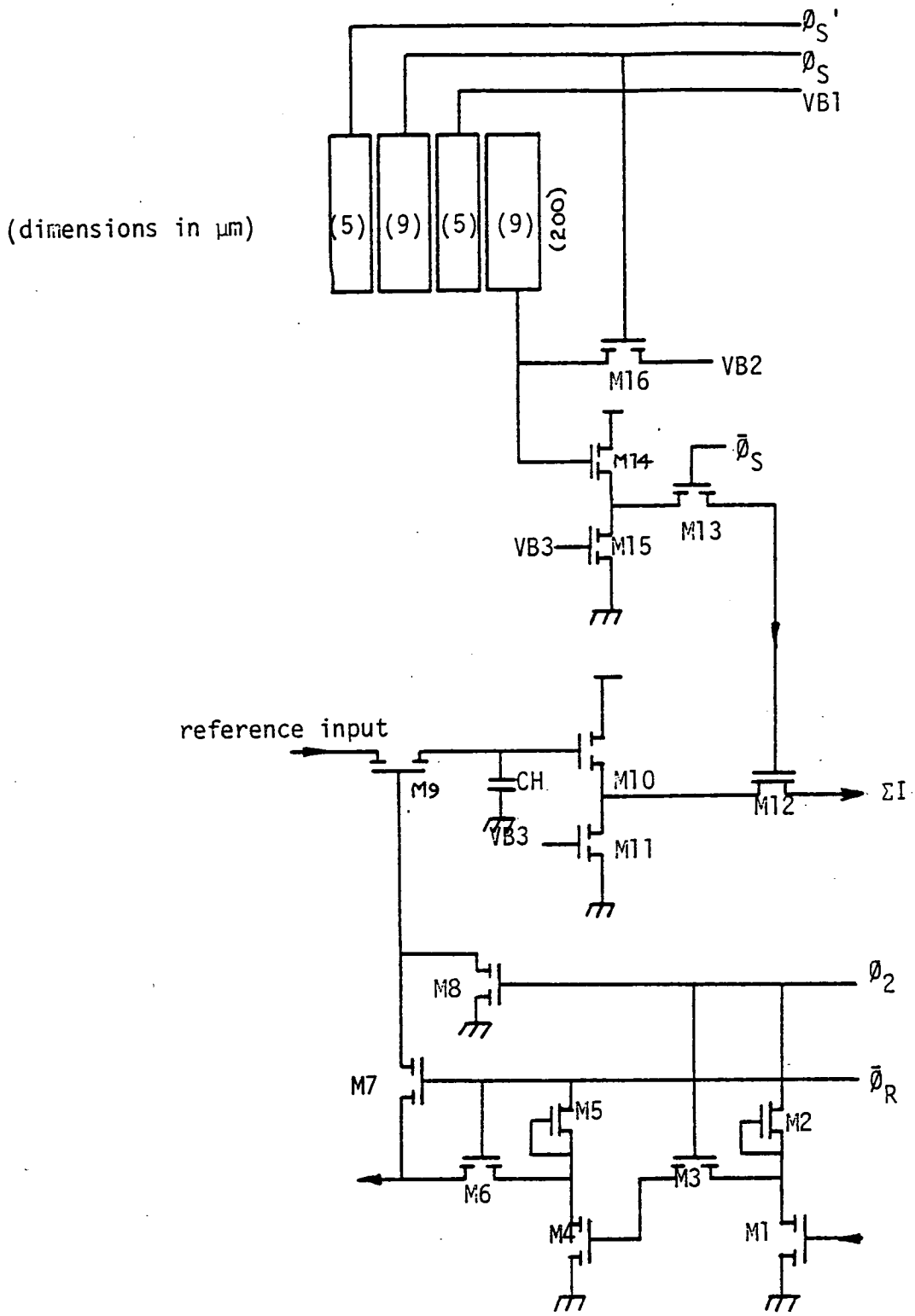


Figure 8.3: PTF II cell circuit schematic

TX	TYPE	L_{DRAWN}	W_{DRAWN}	FINAL AR
M1	P2E	7	18	4
M2	P1D	50	6	1/10
M3	P1E	7	6	1.3
M4	P2E	7	18	4
M5	P1D	50	6	1/10
M6	P1E	7	6	1.3
M7	P2E	8	7	1.3
M8	P1E	7	20	4
M9	P1E	7	6	1.3
M10	P1E	7	70	16
M11	P1E	9	9	1.3
M12	P1D	130	10	1/15
M13	P1E	8	6	1
M14	P1E	7	60	13
M15	P1E	9	9	1.3
M16	P1E	7	6	1.3

Table 8.1: PTF II CELL TRANSISTOR DIMENSIONS

8.3 CELL SPEED AND POWER DISSIPATION

Clearly the power dissipation of the total filter chip is dominated by that of the repeated cell. If the 256-point filter is to maintain a level of dissipation below 300 mW therefore, a design aim in the region of 1 mW per filter point must be achieved. There occurs in the design of every integrated circuit an inevitable trade-off between power-dissipation, speed and layout area. The trade-off is not linear, and a final design must represent only a best compromise, obtained after several iterations.

Consider firstly the tap signal buffers. The design values chosen offer settling to within 1% of expected signal values at a 2 MHz clock rate, with a nominal power dissipation of 300 μW. This speed of reaction is not necessarily important in the tap weight buffers, since the tap weights will normally remain stationary. However these buffers must possess a small-signal output conductance which will not allow serious interference to the applied weight by changes in multiplier current. Here the design values are chosen to limit any tap weight modulaton to 1% or below, as discussed in Chapter 5.

Now consider the reference shift register. The target specification for refresh speed is relatively ill-defined. If refresh from a RAM/DAC system is to be used, then an upper limit of some 100 KHz will be imposed by currently available, moderately priced monolithic DAC's . In any case, if simple CMOS clock generators are to be employed, then they will impose an upper limit of some 5 MHz.

The use of depletion loads offers the best speed-power trade-off for logic gates[31], and bearing in mind the cell power dissipation limit, the device sizes chosen offer 2 MHz operation, using 15V clocks and

dissipating a nominal 375 µW per stage on a 1:1 clock cycle. At lower clock frequencies an assymetrical duty cycle might be used to further reduce this power dissipation.

Thus a total nominal power dissipation of 975 µW per filter cell is anticipated.

8.4 CHIP LAYOUT AND ASSEMBLY

The self-aligned process allows the inclusion of the filter cell within an average pitch of 28µm, detailed in Appendix V. Direct linear repetition of this cell however would lead to an unfavourable chip aspect ratio, so the filter is divided into two blocks, of 128 points each, which are cascaded through the CCD channel by means of a corner-turning diffusion.

This twin-block structure is evident in the photomicrograph of the device shown in Figure 8.4. The remainder of the chip comprises bussing between the two blocks, a small circuit to generate the necessary bias voltages, and two operational amplifiers which control the CCD transfer function. The first amplifier performs feedback linearisation around the CCD input, as described in section 5 (indeed the photomicrograph and results presented there are from this circuit). The second amplifier works on the same clocked-differential-stage principle but is simply connected as a voltage follower to buffer the output from tap 255 for cascading.†

The complete device measures 179 x 149 mils. The latter dimension was critically chosen so that the device might fit within the well of a standard 16-pin DIL package for assembly. Common interconnection of

---

†: Because of the clock cycle delay associated with these amplifiers, it is the penultimate tap which is used to achieve correct timing between cascaded devices.

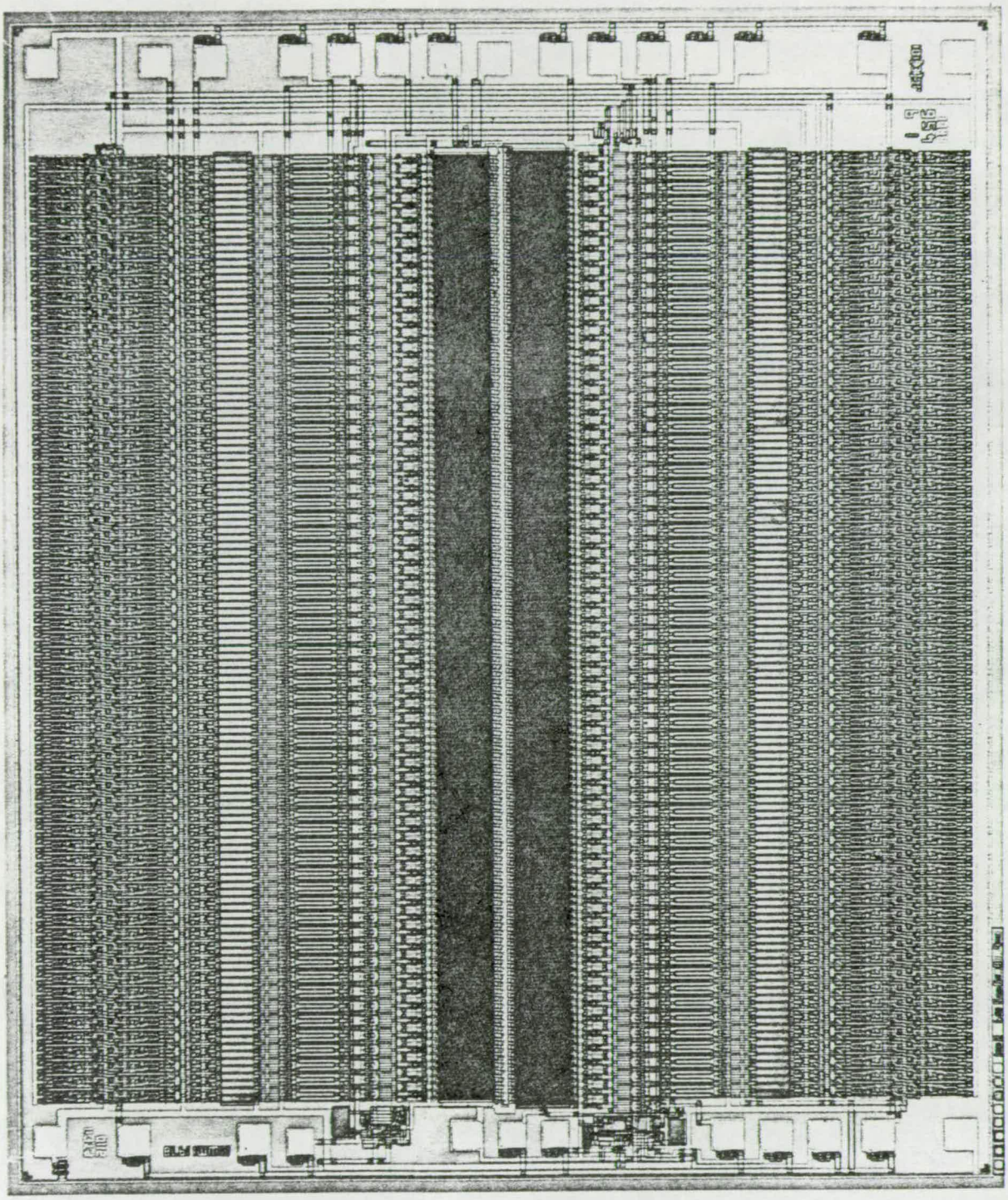


Figure 8.4: Photomicrograph of 256 point PTF II

supplies and clocks on the chip wherever possible merits this reduced pin count, which in turn simplifies the driving requirements.

Three of the pins are used for power supply (+15V, 0V and -5V), two for decoupling the internally generated bias voltages, three for the CCD clocks, four for the reference shift register clocks and signals, and finally four for the analogue signal inputs and outputs.

8.5 DEVICE CHARACTERISATION CIRCUIT

Again it is necessary to drive the device in order to characterise it, but in this case the circuitry, shown in Fig 8.5 is relatively trivial. Above the device is shown the three-phase CCD clock driver in 15V CMOS, along with a similar two-phase driver for the reference shift register. Together these circuits employ two standard CMOS logic packages. Note that drive to the reference clocks is boosted via two bipolar transistors in order to supply current to the shift register gates, through the clocks, in the high state.

Signal input is simply through a d.c.blocking capacitor and bias network. This is the most practical arrangement since the bias level does need to be critically adjusted for the best performance to be obtained from the CCD. Filter output is via a common-base stage and d.c. blocking capacitor. This arrangement is suitable for stationary-reference applications, in keeping with the non-multiplexed 2-quadrant multiplier operation.

The reference waveform is loaded by applying a '1' to the shift register input with the appropriate timing. For device characterisation this signal is supplied from the reference waveform generator. In practice, where the reference might be supplied from a RAM/DAC arrangement

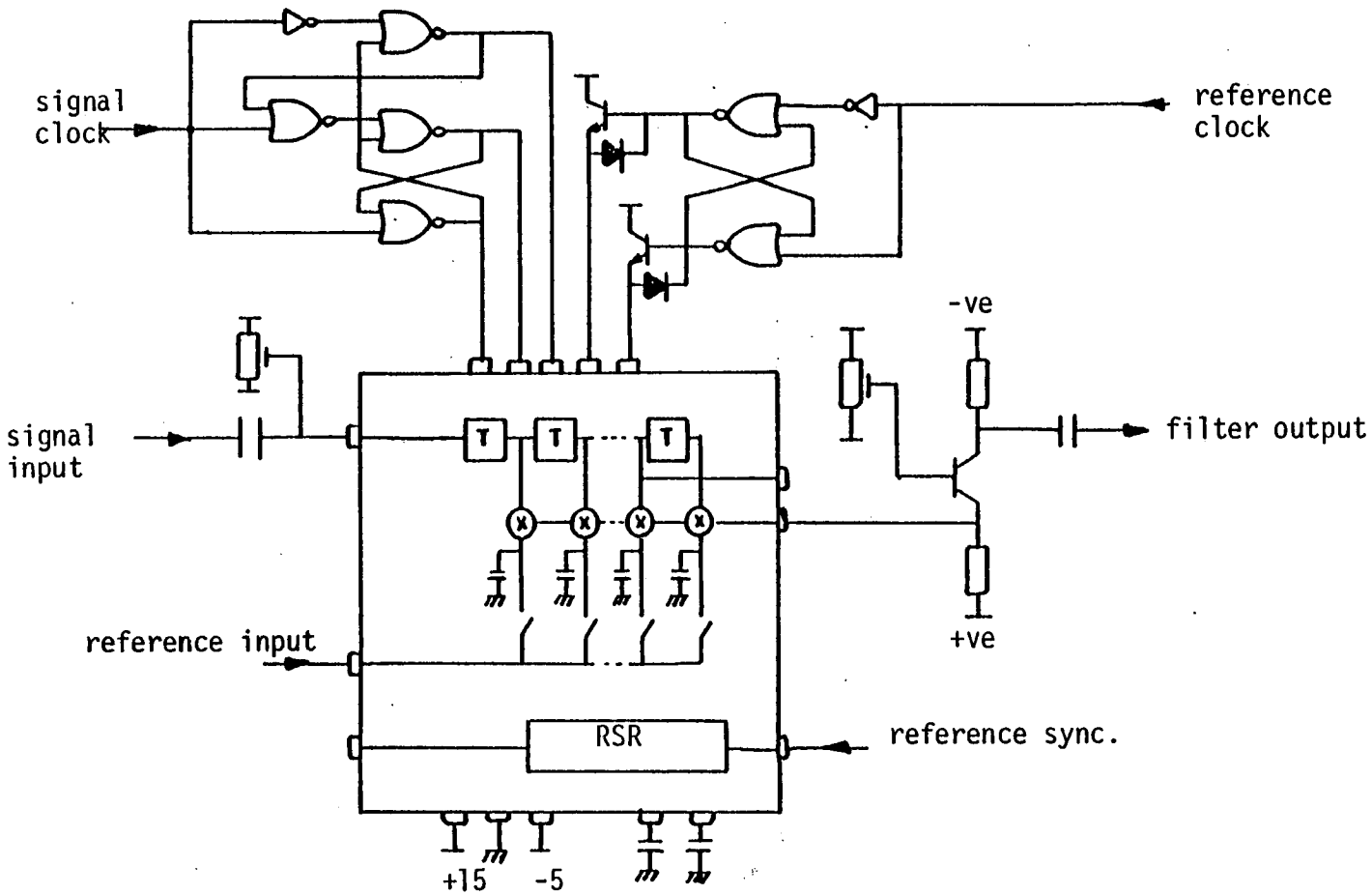


Figure 8.5: PTF II drive circuit

this synchronising pulse is easily derived from the memory address counter.

The remaining device connections are for power supply and decoupling.

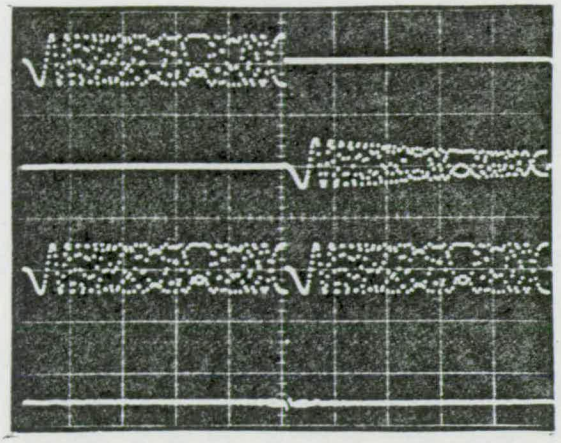
The device has been characterised using the experimental arrangement previously described for the PTF I measurements.

8.6 RESULTS

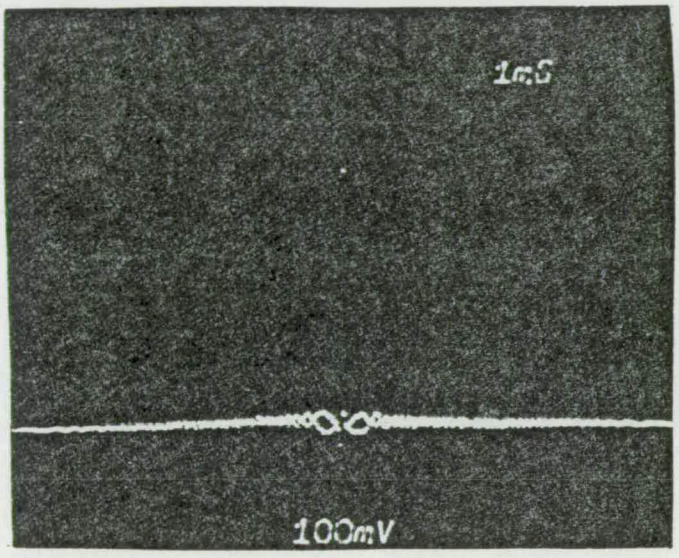
This advanced filter design was also found to be successful, as Fig.8.6 (a) shows. Here a 256-point Nyquist chirp is autocorrelated to produce a single output peak.

Again the quality of the CCD output for cascading, and of the filter output itself are of note. The symmetry of the correlation peak is emphasised in the magnified output trace of Fig.8.6 (b). Charge transfer inefficiency is evident in the delayed signal output, and is here estimated at 10^{-3} per tap. This parameter is largely insensitive to changes in clock and supply levels, but does vary from slice to slice, at least within pilot batches, emphasising the technological dependance of device performance.

Feedback linearisation is an important feature of this filter design in terms of automatic control of the CCD transfer function. Fig.8.7 demonstrates the success of this concept. A 1V p-p signal is shown delayed by 256 samples; the delayed signal is a faithful representation of the input, with a measured loss (before c.t.i. effects) of less than 1% (.09 dB). Only the input bias level, which was set to attain the best CCD performance, was adjusted to achieve this result repeatably between different samples of the device.



(a)



(b)

Figure 8.6: (a) Autocorrelation of chirp 0 - $f_c/2$

(b) Correlation peak detail

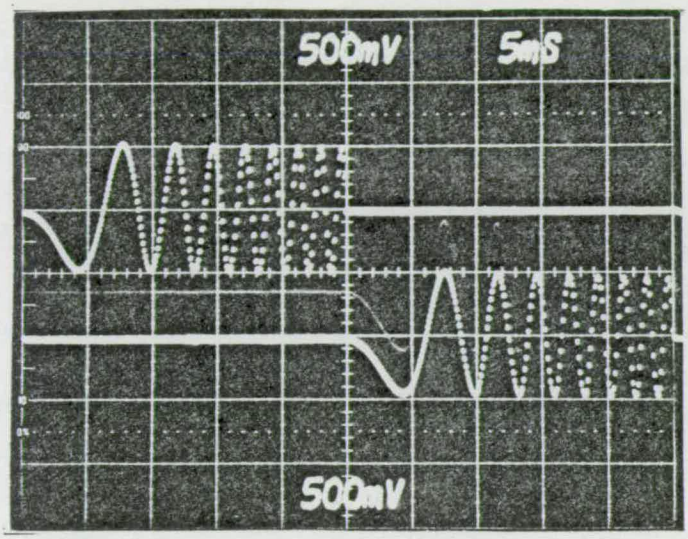


Figure 8.7: Effect of feedback linearisation on delay line characteristic

This result is slightly better than expected, considering the finite open-loop gains of the input and output amplifiers. The discrepancy probably stems from a slight, but favourable, mismatch between the input and output taps because of the small charge injection channel associated with the input. This mismatch increases the input tap load capacitance, causing too much charge to be trapped, and thus compensating for the finite open-loop amplifier gains.

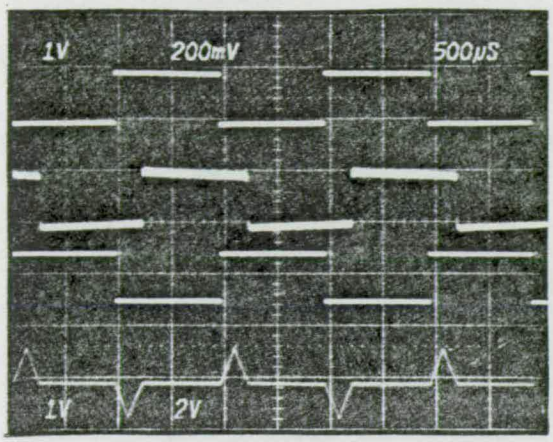
The delay line output is perfectly suitable for direct cascading, as intended, with no further adjustment. Linearity results for this feedbank-linearised delay line have already been presented in Chapter 4, Fig.4.20.

A full test of the bandwidth capability of this device is permitted by the simplicity of the driving circuitry. Fig.8.8 (a) demonstrates square waveform correlation at a signal sample frequency of 1MHz. The sawtooth filter output arises because of the difference in sampling frequency between the signal and reference registers, since the signal clock was here allowed to free run at a much higher rate than the waveform generators. The filter output is correctly maintained above 2MHz sampling, but the CCD output buffer in its present form does not cope with the higher frequency for cascading.

Fig.8.8 (b) demonstrates low frequency, dark-current limited performance at room temperature with a sample clock rate of 500 KHz, corresponding to a total signal delay of $\frac{1}{2}$ second in the CCD. Thermally generated charge biases the CCD into an unacceptably non-linear region for storage times longer than this.

Separate measurements on the tap weight circuitry indicate that refresh rates above 5 MHz are possible because of the fast shift

1 MHz



CCD input

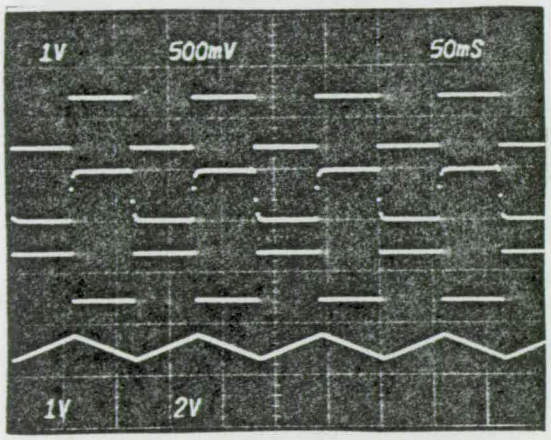
CCD output

reference

filter output

(a)

500 Hz



CCD input

CCD output

reference

filter output

(b)

Figure 8.8: High and low speed operation

(a) f_s = 1 MHz, f_r = 100 kHz

(b) f_s = 500 Hz, f_r = 500 Hz

register action, whilst leakage from the 1 pF storage sites is equivalent to 1V in 20 seconds at room temperature, referred to a typical reference amplitude of 2V p-p.

Thus far the dynamic range of this device has not been considered. The increased number of filter points, and the simplified drive circuitry both suggest an improvement over the 64-point prototype design. This is not the case in practice however. Fig.8.9 shows fixed pattern noise at the filter output, measured at -50 dB rms (-34 dB peak) against a typical correlation peak. This increase in noise level is apparantly due to the simplification of the multiplier operation to 2-quadrant, non-multiplexed mode.

In non-multiplexed mode any common interference, h_i, appearing across the N tap weights, is amplified through the multiplier array as

$$I_i = 2N\beta_m \bar{h} h_i \ , \qquad (8.1)$$

where $\bar{h}$ is the average (rms) tap weight. Thus a small common interference may result in a significant output component, especially if N is large. This effect is otherwise cancelled in multiplexed mode.

Now the fixed pattern noise components in Fig.8.9 are identifiable with the reference refresh waveform, and with the synchronising update impulse. We may deduce that these are weakly (but commonly) coupled to the tap weight holding capacitors and buffers, via capacitive breakthrough and power line interference respectively, and that they therefore become amplified through (8.1).

A return to multiplexed multiplier operation is evidently desirable, especially since system noise present on the reference input

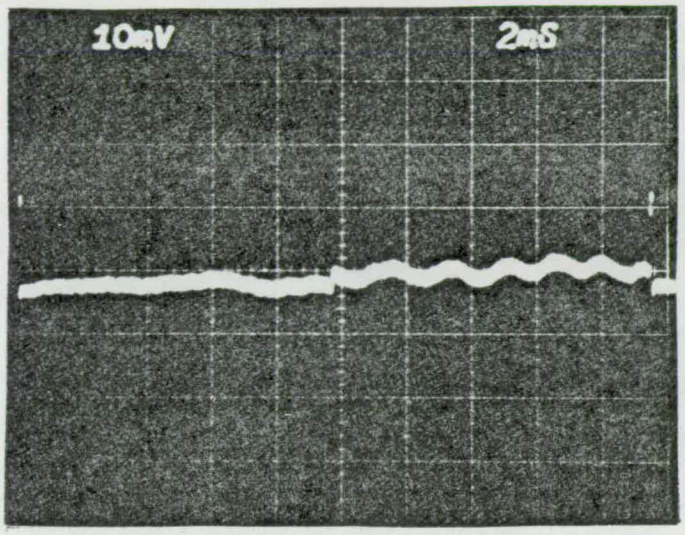


Figure 8.9: Fixed pattern noise

bus in real applications may also reduce the performance of the filter through this effect. This change is being implemented in a proposed redesign.

Table 8.2 summarises the measured performance of the 256-point filter design. Note in particular the reduced cell area, and cell power dissipation, now at 1 mW per point. The CCD dark current and reference decay parameters both reflect process improvements, whilst the surface-channel charge transfer inefficiency remains at 10^{-3} per tap.

These results support the feasibility of a practical high time-bandwidth programmable matched filter for compact sonar. An application programme is currently in progress, using two cascaded devices plus digital reference memory, to form a 'black-box' 512-point filter module on a single Eurocard.

Figure 8.10 demonstrates the successful operation of this module as a matched filter. In the upper photograph, the matched filter response to a 512-point up-chirp is shown, exhibiting the expected correlation peak. The lower photograph shows the filter response to the same input signal contaminated by white noise. Note that the correlation peak is still easily detectable above the noise, through the processing-gain effect. An improvement ratio of 20dB is estimated. It is interesting to compare this with the theoretical processing gain predicted by (1.3) as;

$$10 \log_{10} N/2 = 24 \text{ dB} \qquad (8.2)$$

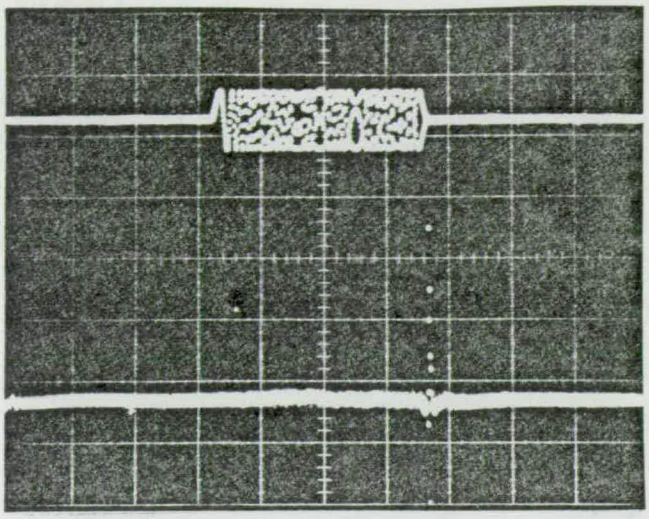
The 4 dB loss in processing gain may be associated with a reduction in filter sensitivity due to charge transfer inefficiency, as treated in Appendix II. However, Buss[98] has shown that the input noise also

Number of Filter Points		256
Power Dissipation		1 mW/point
Cell Area		1.1 x 56 mils/point
Charge Transfer Inefficiency		10^{-3} per point
Harmonic Distortion	(1)	-40 dB
CCD Through-gain		> .99
Clock Rate	(2)	500 Hz - 1 MHz
Dynamic Range	(3)	50 dB
(up-chirp correlation)	(4)	60 dB
Dark Current	(2)	5 nA/cm^{-2}
Reference Decay	(2)	0.05 V sec^{-1}

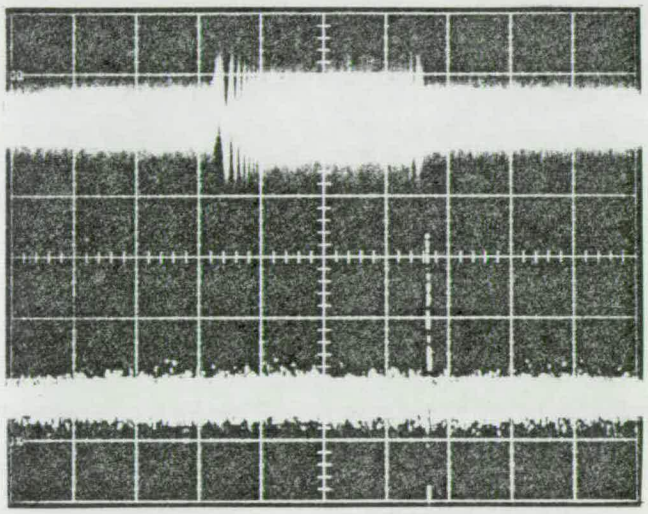
Notes:
1. Signal 1v p-p, reference $\pm$1v
2. Room temperature
3. With FPN
4. Without FPN

Table 8.2: PTF II Performance Summary

suffers a sensivity reduction through charge transfer inefficiency, to the extent that the processing-gain is not,to first order, sensitive to this parameter. Its effect here is attributed to a gross inefficiency ($\simeq 5 \times 10^{-3}$ per point) present in the two particular devices used for the experiment. Despite this, the result is a convincing demonstration of noise rejection through the matched filter module.



(a)



(b)

Figure 8.10: 512 point matched filter operation

(a) matched filter response to chirp from 0 - fc/4 Hz.

(b) matched filter response to noise contaminated chirp.

Scales: Horizontal 10 msec/cm

Input chirp 2V/cm

Filter output 12V/cm

f_c = 20 kHz

CHAPTER NINE : LIMITATION ANALYSIS

The design and operation of a particular form of programmable transversal filter have been reported. It is of some interest to evaluate the predictable effects of errors and noise within the structure in the light of these results.

9.1 TAP WEIGHT ERROR

Static operational 'defects' of these integrated structures include non-linearity, non-uniform d.c. offsets and non-uniform gains in both the signal and reference channels. From all of these effects the contribution to the correlation from tap n may be expressed quite generally as;

$$(1 + a_n)\ (x_n + b_n + c_n x_n + d_n)(h_n + e_n + f_n h_n + g_n) \qquad (9.1)$$

where x and h are the desired tap signal and weight values, and each of the error coefficients is small and may be attributed a physical significance:

- a_n - multiplier gain variation
- b_n - signal offset
- c_n - tap gain variation
- d_n - signal distortion polynomial
- e_n - reference offset
- f_n - reference gain variation
- g_n - reference distortion polynomial

Equation (9.1) may be rearranged as;

$$(m_n + s_n + i_n)(j_n + r_n + k_n r_n + l_n) \qquad (9.2)$$

where $m_n = b_n(1 + c_n)^{-1} \simeq b_n$

$i_n = d_n(1 + c_n)^{-1} \simeq d_n$

$j_n = e_n(1 + a_n)(1 + c_n) \simeq e_n$

$k_n = a_n + c_n + a_n c_n + f_n(1 + a_n)(1 + c_n) \simeq a_n + c_n + f_n$

$l_n = g_n(1 + a_n)(1 + c_n) \simeq g_n.$

Given that signal offsets are cancelled via the multiplication algorithm, it is interesting to note that all of the errors, apart from signal distortion, may be written into the weighting factor. Since the reference is programmable and stationary, it could be predistorted to correct for all weighting coefficient errors. Indeed, it should be possible to montior the impulse response of the filter and iteratively correct the reference until an arbitrarily accurate response is obtained.[61]

Returning to the uncorrected case, and ignoring signal offsets n_n, convolution using equation (9.2) yields

$$\underline{x} * \underline{h} + \underline{x} * (\underline{j} + \underline{kh} + \underline{l}) + \underline{i} * (\underline{j} + \underline{kh} + \underline{l}) \qquad (9.3)$$

where * denotes convolution and - denotes a vector, and the distortion polynomials i and l are assumed to be positionally invariant.

Clearly, in addition to the expected correlation, between $\underline{x}$ and $\underline{h}$, various cross correlation components exist. The effect of these components upon filter performance depends upon the statistical

properties of the error vectors and of the signals, and upon the mode of use. The error vector $\underline{j}$ represents threshold offset variations in the reference buffers and will be predominantly random. However $\underline{k}$ is a combination of the tap and multiplier gain errors, which may will exhibit a spatial coherence as oxide thickness changes slowly over a chip. More significantly, $\underline{k}$ may contain step functions where physical discontiniuties exist, such as between the two blocks within the 256-point filter, or (especially) between cascaded devices.

Consider a chirp matched filtering application. The chirp waveform has a 'thumb-track' ambiguity diagram. It is often chosen for detection systems because of its ability to reject interference. In this context both the $\underline{j}$ and $\underline{k}$ vectors present in these filters are rejected as 'noise'. As the number of points in the filter is increased so is the processing gain, so that long matched filters become even more resistant to device defects.

A more significant error can occur through correlation of the distorted signal and reference vectors. Distortion increases frequency components in the waveform, having an effect similar to a severe Doppler shift (reflection from a moving target). Since the distortion is composed of sinusoids, these can correlate with parts of the original chirp. Fig.9.1 illustrates this effect for second harmonic distortion on a Nyquist-bandwidth chirp. The exagerrated distortion component is a chirp from $0\text{-}2f_n$ Hz, which with aliasing appears as a V-chirp from $0\text{-}f_n\text{-}0$ Hz. Correlation of this V-chirp with the undistorted reference gives a chirp output pattern over the duration of the correlation , which may interfere with the natural sidelobe pattern. This is not particularly significant for unweighted chirp filtering, where the sidelobes remain significant in any case,

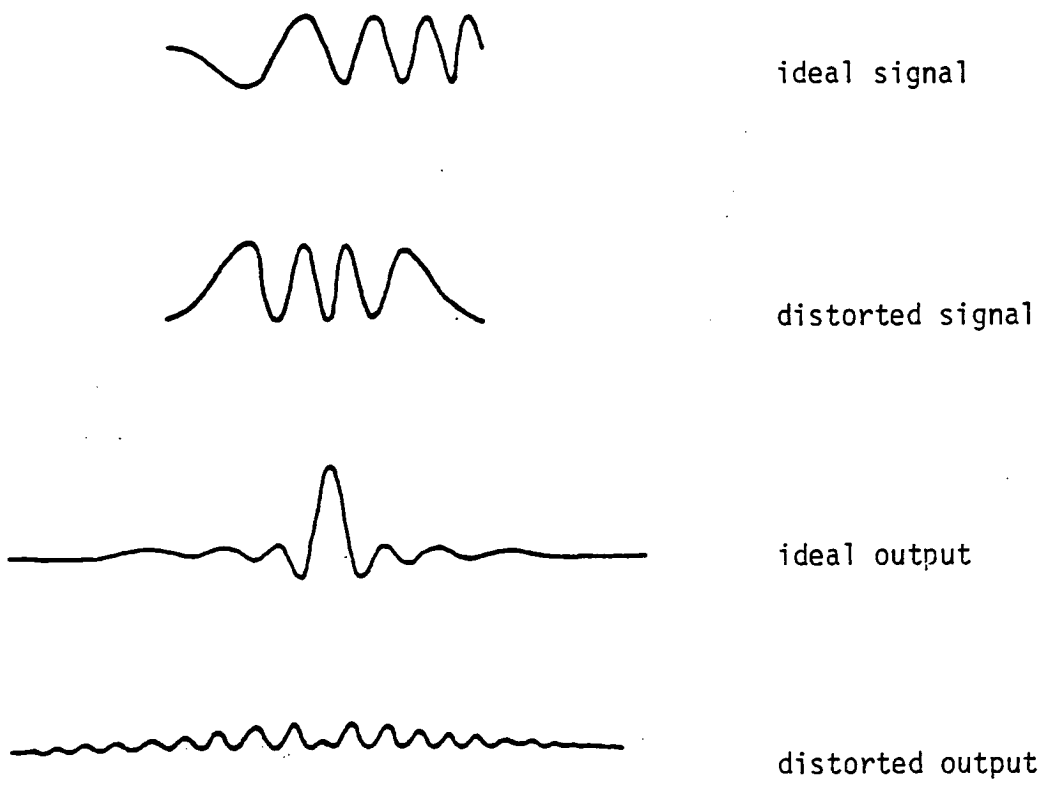


Figure 9.1: Effect of Signal Distortion on Matched Filter Performance

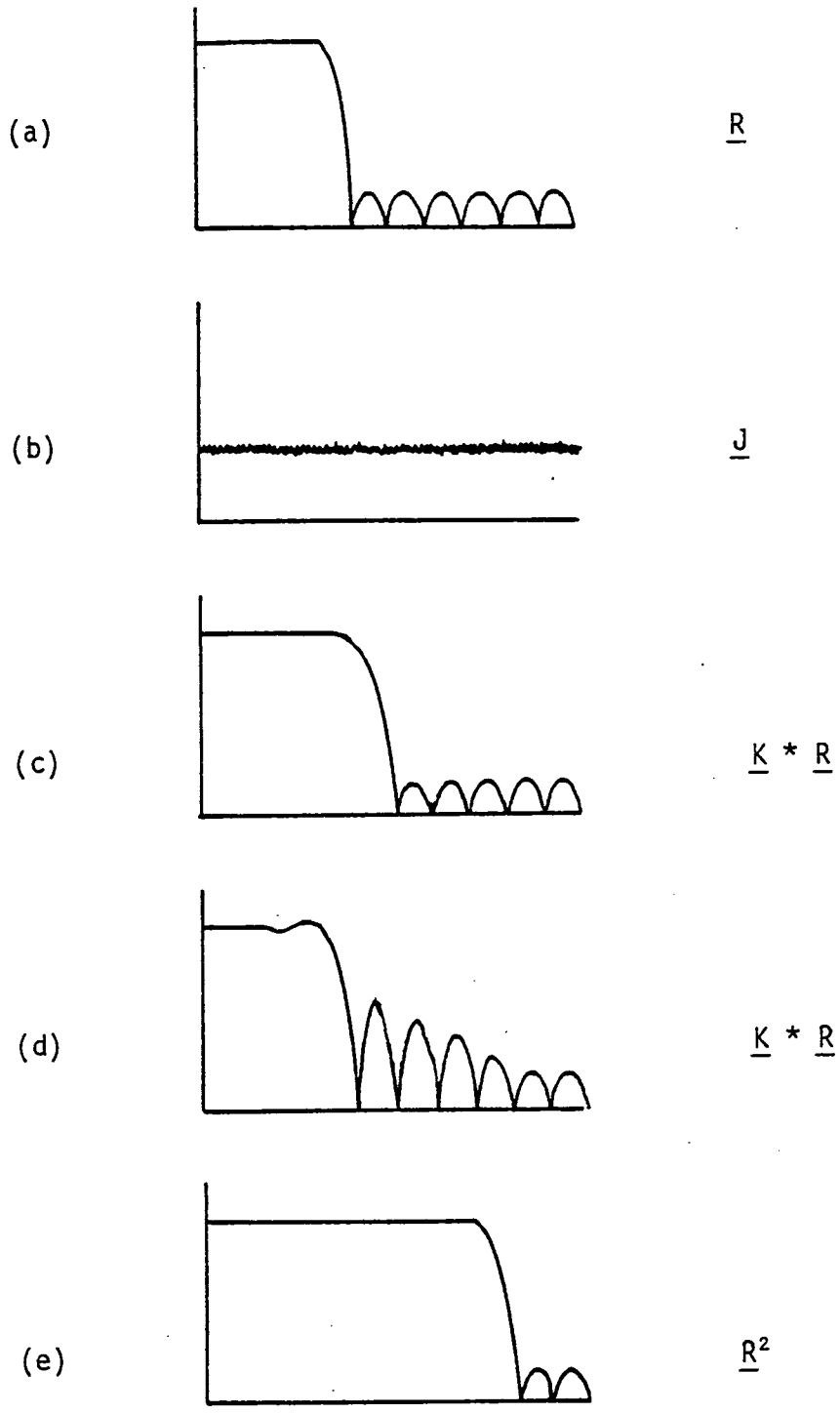


Figure 9.2: Effects of filter errors on low-pass characteristic

(a) ideal
(b) random reference offset
(c) spatially coherent low-frequency tap gain error
(d) step function tap gain error
(e) reference distortion

but where weighting is used for sidelobe suppresion this interference could limit performance. The effect is independent of the number of filter points used.

Consider now a frequency filtering application, The random error vector $\underline{j}$ has a flat spectrum correlating with all frequency components. This will mar the stop-band performance of any filter, according to the variance of the error distribution[60]. Coherent error components in $\underline{k}$ can also affect the stopband. This effect is more complex, since the signal is actually correlated against $\underline{kr}$, having a frequency distribution which is the convolution of these two vectors in the frequency domain.

Fig.9.2 shows the approximate form of these error responses. Low spatial frequency gain variations produce an error response similar to, but slightly broader than, the main pass-band response. Step function errors in $\underline{k}$ result in Gibbs ripples at the pass-band edge. Reference distortion results in a broad pass-band error characteristic.

Clearly the complete filter characteristic is a combination of all of those shown in Fig.9.2, with varying degrees of emphasis. Reference to the real low-pass filter example of Fig.7.9 would suggest the dominant presence of a random offset error $\underline{j}$, and a step function in $\underline{k}$. This step function probably arises through a general reference offset, caused by incorrect adjustment of the reference zero level. This is an unavoidable failing of this type of filter structure as no automatic method of setting this level has yet been devised.

9.2 NOISE

CCDs and MOSTs are essentially low noise devices, and in a transversal filter this noise is reduced further by the processing gain. In

this section the effect of the major noise sources within the filter is presented. These sources are grouped according to whether they occur in the signal or the reference channel, or directly at the filter output.

Five noise sources are present in the signal channel:

- input noise
- dark current noise
- fast interface state noise
- tap reset noise
- buffer noise

Now any noise source in the signal channel will be filtered to some extent through the multipliers. The r.m.s. output noise voltage is given by a modification of the convolution equation:[81]

$$v_o = R_s \beta_m \sum_{n=0}^{N-1} \sum_{m=0}^{N-1} h_n h_m \langle x_n x_m \rangle^{\frac{1}{2}} \qquad (9.4)$$

where $\langle x_n x_m \rangle$ is the covariance of the noise source. Evidently the output noise voltage is a function of both the filter weighting coefficients, and the statistics of the various noise sources.

Input noise is composed of uncertainty in the reset level of the input tap, and also in setting the CCD surface potential. With on-chip feedback linearisation there may be a further source associated with reseting the operational amplifier differentiating capacitor; other noise sources in the amplifier are attenuated by the value of the open-loop gain. Assuming these sources are characterized by kTC noise, each has a covariance[39],

$$\begin{aligned} \langle x_n x_m \rangle &= \left(\frac{kT}{C}\right) \quad , n=m \\ &= 0 \quad , n \neq m. \end{aligned} \qquad (9.5)$$

Inserting (9.5) into (9.4) gives

$$(v_o)_{in} = R_s\beta_m\left(\sum_{n=0}^{N-1} h^2 n\right)^{\frac{1}{2}} \left(\frac{kT}{C}\right)^{\frac{1}{2}}. \tag{9.6}$$

Intermediate reset noise sources in the CCD channel, at corner-turns and between cascaded devices, are characterized in the same way,

$$(v_o)_{int} = R_s\beta_m\left(\sum_{n=\ell}^{N-1} h^2 n\right)^{\frac{1}{2}} \left(\frac{kT}{C}\right)^{\frac{1}{2}}, \tag{9.7}$$

where ℓ is the tap position immediately following the source.

The effects of thermal noise can be important if the device is to be operated at high temperatures or low frequency. The covariance function for dark current noise is[39],

$$\begin{aligned} \langle x_n x_m \rangle &= n\left(\frac{qJ_D A}{C_L{}^2 f_c}\right) \quad , n=m \\ &= 0 \qquad\qquad\quad , n\neq m \end{aligned} \tag{9.8}$$

where J_D is the average dark current density, A is the CCD cell area, and C_L is the tap load capacitance. This gives an output noise voltage of

$$(v_o)_{th} = R_s\beta_m\left(\sum_{n=0}^{N-1} nh_n{}^2\right)^{\frac{1}{2}} \left(\frac{qJ_D A}{C_L{}^2 f_c}\right) \tag{9.9}$$

Trapping noise due to fast surface states is correlated between adjacent charge packets and has a covariance function of the form,

$$\begin{aligned} \langle x_n x_m \rangle &= 2n\nu_{fis}^2 \qquad\qquad ,n=m \\ &= -n\nu_{fis}^2 \qquad\qquad ,n=m-1,\ n=m+1 \\ &= 0 \qquad\qquad\qquad , \text{otherwise}. \end{aligned} \tag{9.10}$$

where

$$\nu^2_{fis} = \frac{q^2 A N_{ss} k T \ln 2}{C_L^2} \quad . \tag{9.11}$$

The output noise voltage due to fast interface states is thus,

$$(\nu_o)_{fis} = R_s \beta_m \nu_{fis} \sum_{n=0}^{N-1} n(2h_n - h_n h_{n-1} - h_n h_{n+1})^{\frac{1}{2}} . \tag{9.12}$$

Tap reset noise is statistically independent, having a covariance function and output noise voltage as in equations (9.5) and (9.6).

Thermal noise in the MOST tap buffers stems largely from the active device in these configurations and has a covariance

$$\begin{aligned} \langle x_n x_m \rangle &= \frac{6kTB}{g_m} \quad , n=m \\ &= 0 \quad , n \neq m \end{aligned} \tag{9.13}$$

where B is the system bandwidth. The output noise voltage becomes

$$(\nu_o)_{buff x} = R_s \beta_m \left(\frac{6kTB}{g_m}\right)^{\frac{1}{2}} \left(\sum_{n=0}^{N-1} h_n^2\right)^{\frac{1}{2}} \tag{9.14}$$

Note that where the multiplier is time multiplexed, and the signal zero is switched in before the tap buffer, the resulting double-correlated sampling effect virtually eliminates this noise source.

Noise sources in the reference channel are similarly eliminated by the differential multiplier action, they then appear not as random noise at the output, but rather as time-variant errors in the tap weight values. Where the multiplier is not multiplexed however, reference noise sources ν_n pass directly to the filter output through the

multiplier transistor channels according to,

$$v_o = v_n g_{ds} R_s \simeq v_n R_s \beta_m (V_{GS_o} - V_T). \tag{9.15}$$

Generally only one tap weight is reset in any one sample period so,

$$(v_o)_{hrst} = R_s \beta_m (V_{GS_o} - V_T) \left(\frac{kT}{C_h}\right)^{\frac{1}{2}}. \tag{9.16}$$

Thermal noise is generated in all tap weight buffers however,

$$(v_o)_{h\ buff} = R_s \beta_m (V_{GS_o} - V_T) N^{\frac{1}{2}} \left(\frac{6kTB}{g_m}\right)^{\frac{1}{2}}. \tag{9.17}$$

Thermal noise generated in the multiplier channels adds directly into the output;

$$(v_o)_{mult} \simeq R_s (4kTB\ N \beta_m (V_{GS_o} - V_T)^{\frac{1}{2}}. \tag{9.18}$$

Finally, thermal noise in a common base bipolar summing stage is given by;

$$(v_o)_{sum} = 4kTB(R_S + R_B)^{\frac{1}{2}}. \tag{9.19}$$

The peak signal amplitude at the output of an unweighted matched filter is;

$$(v_o)sig = R_s \beta_m \sum_{n=o}^{N-1} h_n^2. \tag{9.20}$$

This estimate is strictly in error because charge transfer inefficiency causes signal loss. It has been computed[84] that for Nyguist rate up-chirps the peak loss is of the order of 1dB for $N\varepsilon = .25$, rising to some 9dB for $N\varepsilon = 2$.

These noise and signal expressions are evaluated for the 64-point and 256-point devices, and for a hypothetical 2048 point filter based on cascaded 256-point devices, using ideal up-chirps of $BT_d = N/2$ in all cases. Table 9.1 lists the parameters used and the results are tabulated in Table 9.2.

Summing resistors have been chosen to produce correlation peaks of similar amplitude before c.t.i. loss, and the filters are compared over a typical sonar bandwidth of 25kHz (f_c = 50kHz). The only major difference in the parameters is a dark current improvement associated with the polysilicon gate process.

Data for the 2048-point filter assume a return to time-multiplexed operation. This result is intended to illustrate the potential dynamic range available in a long matched filter application. Buss et al[54] have concluded that processing gain is not effectively enhanced for these filters above $TB\varepsilon = 1$, which corresponds to 2000 points at $\varepsilon = 10^{-3}$ per tap.

The results show clearly that dark current and surface state noise dominate the dynamic range in these applications. The two effects are approximately equal for the 64 and 256 point devices, though the result for the latter device would have been worse but for the improved dark current parameter. Calculations therefore predict dynamic ranges around 75 dB for both devices when applied as up-chirp matched filters. The measured value of 64 dB for the 64-point device agrees reasonably

PARAMETER	64 point	256 point	2048 point
R_S, R_B	18kΩ	4.7kΩ	560Ω
β_M	1μA/v^2	1μA/v^2	1μA/v^2
B	25 kHz	25 kHz	25 kHz
C_O	0.3 pF	0.33 pF	0.33 pF
C_L	0.3 pF	0.33 pF	0.33 pF
C_{opamp}	-	1 pF	1 pF
C_{chan}	-	0.5 pF	0.5 pF
J_D	20nA/cm^2	5nA/cm^2	5nA/cm^2
A	6720μm^2	5600μm^2	5600μm^2
f_c	50 kHz	50 kHz	50 kHz
N_{SS}	10^{10}cm^{-2}eV^{-1}	10^{10}cm^{-2}eV^{-1}	10^{10}cm^{-2}eV^{-1}
g_m	100μS	100μS	100μS
$V_{GS_o} - V_T$	4v	10v	10v
$\hat{x}$	.5v	.5v	.5v
$\hat{h}$	1v	1v	1v

Table 9.1: Parameters for Evaluation of Noise Expressions

Noise Source	64 point	256 point	2048 point
Diode cut-off input	12μV	6μV	2μV
Input tap reset	12μV	6μV	2μV
Input op amp	-	4μV	1μV
Intra-channel reset	-	4μV	10μV
Dark current	33μV	30μV	74μV
Surface states	37μV	32μV	30μV
Tap reset	12μV	6μV	2μV
Tap buffer	.4μV	.1μV	.04μV
Weight reset	-	3μV	-
Weight buffer	-	2μV	-
Multiplier	8μV	.5μV	2μV
Summing	.6μV	2μV	.7μV
Total r.m.s. noise	55μV	46μV	81μV
Peak signal (less noise)	275 mV	268mV	102mV
Dynamic range	74 dB	75 dB	62 dB

Table 9.2: Calculated Dynamic Range for 64 and 256 Point Matched Filters and for Hypothetical 2048 Point Filter

well. The additional measured noise might be attributed to errors in the estimated values of N_{SS} and J_D. The same agreement is offered by the measured results for the 256 point filter, if the spurious fixed pattern noise is neglected. Alternatively the additional noise may in either case be induced by the driving circuitry, rather than by any further internal sources.

The extension of these calculations to the hypothetical 2048-point cascaded filter is of considerable interest. Most of the noise sources, apart from the dominant ones, become less significant, because of the 'processing-gain' effect. Among these sources however, intra-channel reset noise has increased because of the multiple input-output structures within the cascaded chain. Of the two dominant sources, dark current noise has become far more significant, and this is because of an interesting property of the up-chirp.

Surface state noise is cumulative and adopts a frequency dependent transfer function of the form given in (3.8). As the noise accumulates along the line the lower frequency components tend to be cancelled, whereas the higher frequencies are accentuated. Now for up-chirp correlation, the reference consists only of lower frequencies at the higher tap numbers, and the surface state noise is therefore cancelled to some extent.

This is not true of dark current noise, which is effectively white and accumulates as $N^{1.5}$, whereas the signal accumulates only as N. For this reason, the dynamic range falls as the number of points is increased. Correlation peak loss compounds this problem, so that a lowered dynamic range of 62 dB is expected for the 2048-point filter.

This is probably the lowest tolerable dynamic rang for practical device application, and thus reinforces the BT_d limit of 1000 for CCD matched filters.

9.3 FILTER LENGTH LIMITATION DUE TO CHARGE TRANSFER INEFFICIENCY

Large time-bandwidth products are desirable in matched-filter detectors if processing gain is to be maximised. However, charge transfer inefficiency progressively corrupts the signal information and thus the form of the correlation peak, and must ultimately impose a limit on the maximum filter length that may be used. In particular, c.t.i. may be expected to cause a peak positional error (time shift), and a degradation in peak sharpness and in peak-sidelobe ratio.

The effect of charge transfer inefficiency on the performance of the chirp detection system that was shown in Figure 1.5 is studied here by computer simulation. The linear c.t.i. model, proposed by Vanstone[25], is used.

Many simulations of the system have revealed two interesting factors which enable the results to be simplified and generalised. Firstly, for a given system, c.t.i. effects for a chirp of given BT_d depend *only* upon the number of points used to realise the filter. Thus increasing the number of filter points N (above the minima derived in section 1) for a given ε does not affect the form of the correlation peak, despite the inherent increase in cumulative transfer inefficiency $N\varepsilon$. This is because any increase in the filter length used for a given chirp must increase the relative sampling frequency, reducing c.t.i. effects (because of over-sampling) by approximately the same factor.

Secondly, for each system, the normalised form (or shape) of the correlation peak is similar for all combinations of the product εBT_d having the same value, regardless of the values of the component

variables, or of the number of filter points. Quantitatively, we may expect this result if we consider two matched filters of length N, operating at the same sampling frequency. If that filter matching a chirp of higher bandwidth has a corresponding lower value of ε, then we may expect c.t.i. effects to be similar.

These results enable the system to be characterised by a single set of normalised peaks, corresponding to different εBT_d products. The normalised results are shown in Figure 9.3. Note the progressive degradation of the correlation peak. The critical point, beyond which the filter is deemed to be unusable, must depend upon the particular application. However, for the sonar matched filter application, a suitable criterion appears to be

$$\varepsilon BT_d < 1.0 \qquad (9.21)$$

With ε at 10^{-3} per point, CCD matched filters may be employed for time-bandwidth products not exceeding 1000.

Buss[98] has noted that additive noise at the filter input is also attenuated by charge transfer inefficiency, so that the processing-gain of the filter is , to first order, insensitive to this effect. Indeed, the expected loss in processing gain for $\varepsilon BT_d < 1.0$ is less than 3dB.

These results support those presented in the previous section on noise analysis, both suggesting a technology-limited restriction of some 2000 points for the CCD matched filter.

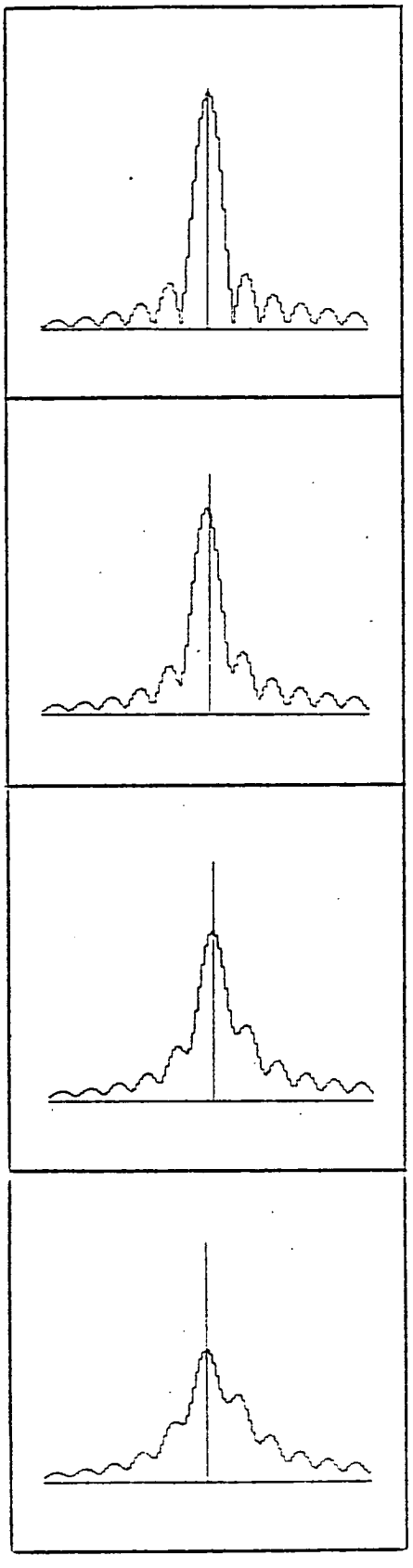


Figure 9.3: Normalised correlation peak degradation for increasing εBT_d.

CHAPTER TEN : CONCLUSIONS

A direct form architecture has been adopted for the implementation of compact programmable transversal filters, for application in light-weight, low-power sonar equipment. CCD technology has been used to form a simple, high-density tapped delay line realisation, and analogue tap weight storage and multiplication have been implemented using advanced linear MOS techniques.

The total CCD transfer function may be automatically controlled by the technique of feedback linearisation. A novel operational amplifier, monolithically compatible with the CCD has been developed, and a fully integrated structure has been successfully demonstrated at a 1 MHz clock rate. This facility allows CCD delay lines (and filters) to be directly cascaded with a minimum of ancillary circuitry, and guarantees a uniform linear characteristic free from insertion loss.

The difficulty of achieving stable analogue multiplication over a wide dynamic range has previously hampered filter development. A stable, time-multiplexed, single transistor 4-quadrant multiplier has been developed here which is;

- linear and accurate to 1%
- free from mismatching errors
- free from drift
- optimally compact

It is used to maximum effect in two practical filter realisations.

These developments have lead to the successful realisation of cascadable 64-point and 256-point programmable filters, which find major application in large time bandwidth, low power sonar systems. A potential bandwidth of 1 MHz exceeds the sonar requirement, and so other applications, in medical ultrasonics and high speed data transmission might be considered.

Error analysis shows that matched filtering applications are generally tolerant to all of the circuit imperfections present in these filters, especially where large numbers of points are used for high time-bandwidth products. If specially weighted references are used for sidelobe suppresion however, good multiplier linearity is essential.

Examination of the frequency filtering problem has shown a general sensitivity to error components. In particular, care must be taken to avoid sharp variations in gain between filter points, by using only single, linear arrays, and avoiding cascading. If these devices are to be used in demanding frequency filtering applications, then further attention should be given to the iterative reference correction techniques, since these are capable of automatically correcting for errors to yield arbitrarily defined, but well controlled, filter characteristics.

Noise calculations have been presented which support measurements made on the two devices. Dark current noise emerges as the limiting factor in long filter realisations, such that the time-bandwidth product should not exceed approximately 1000, if a dynamic range of 60 dB is to be maintained.

A study of the effect of charge transfer inefficiency on the form of the correction peak supports the time-bandwidth limit of 1000 for

the surface-channel CCD matched filter.

Further development of these filters is progressing. The 256 point design is being refined to eliminate a fixed pattern noise problem, and a new 64-point design is in commission for short filter applications. Other filter structures based on these principles are also in development, most notably a 64-point monolithic adaptive filter, in which the tap weights are internally computed to yield on optimised filter response.

Digital filter realisations are always advancing, especially so with the evolution of fast multipliers. Whether these will eventually obviate the analogue approach is not in doubt, as VLSI technology will surely make possible dense digital filters of equivalent processing power, but without the inevitable vagaries of analogue circuitry. Until this technology becomes available however, the CCD based filters reported here enjoy an unquestionable power, size and weight advantage. It is likely that the superior bandwidth offered by parallel processing in the analogue devices will indefinitely maintain their advantage at least for higher frequency application.

As a final testament to the effectiveness of these devices, Figure 10.1 shows a high resolution sonar return obtained from real-time trials in shallow water, using the prototype 64-point device in the sonar system of Figure 1.5, with a BT_d of 128. A target group is distinctly visible above the background reverberation. The CCD system performance was indistinguishable from a benchmark digital system.

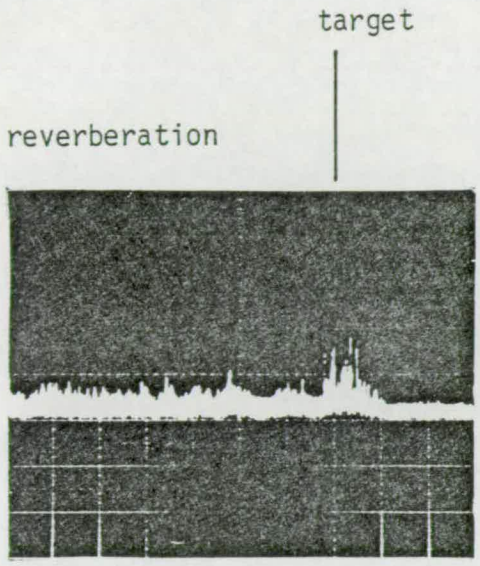


Figure 10.1: High resolution sonar return from an underwater target in shallow water

REFERENCES

1. Boyle, W S and Smith, G E, "Charge-coupled Semiconductor Devices", Bell Syst Tech J, 1970, 49, pp 587-593.

2. Sze, S M, "Physics of Semiconductor Devices", John Wiley and Sons, 1969.

3. Grove, A S, "Physics and Technology of Semiconductor Devices", John Wiley and Sons, 1967.

4. Cobbold, R S C, "Theory and Applications of Field-Effect Transistors", John Wiley and Sons, 1970.

5. Gosling, W, Townsend, W G and Watson,J, "Field Effect Electronics", Butterworth and Co, 1971.

6. Crawford, R H, "MOSFET in Circuit Design", McGraw-Hill, 1967.

7. Sequin, C H and Tompsett, M F, "Charge Transfer Devices", Academic Press, New York, 1975.

8. Hobson, G S, "Charge Transfer Devices", Arnold, 1978.

9. Howes, M H and Morgan, D V, "Charge Coupled Devices and Systems", Wiley, 1979.

10. Beynon, J D E and Lamb, D R, "Charge Coupled Devices and their Applications", McGraw-Hill, 1980.

11. MacLennan, D J, "Design and Operation of a Programmable Charge-Coupled Device Transversal Filter", Thesis, University of Edinburgh, 1976.

12. Sequin, C H, Mohsen, A M, "Linearity of Electrical Charge Injection into Charge-Coupled Devices", IEEE J Solid-State Circuits, 1975, SC-10, pp 81-92.

13. Wallinga, H, "A Comparison of CCD Analogue Input Circuit Characteristics", Proc of Int Conf on Technology and Applications of CCDs, Edinburgh, September 1974, pp 13-21.

14. Tompsett, M F, "Surface Potential Equilibration method of setting charge in charge-coupled devices", IEEE Trans. Electron Devices, 1975, ED-22, pp 305-309.

15. McCaughan, D V and Harp J G, "Phase Referred Input: A simple new linear CCD input method", Electron. Lett., 1976, 12, pp 682-683.

16. MacLennan, D J and Mavor J, "Novel Technique for the Linearisation of Charge-Coupled Devices", Electron Lett, 1975, 11, pp 222-223.

17. MacLennan, D J and Mavor J, "Linearisation of the Charge Coupled Device Transfer Function", Proc of CCD Applications Conf, San Diego, October 1975, pp 291-294.

18. Carnes, J E, Kosonocky, W F and Ramberg, E G, "Drift-aiding Fringing Fields in Charge-Coupled Devices", IEEE J. Solid State Circuits, 1971, SC-6, pp 322-326.

19. Ibid, "Free Charge Transfer in Charge-Coupled Devices", IEEE Trans, Electron Devices, 1972, ED-19, pp 798-808.

20. Strain, R J and Schryer, N L, "A Non-Linear Diffusion Analysis of Charge-Coupled Device Transfer", Bell Syst Tech J, 1971, 50, pp 1721-1740.

21. Kim, C and Lenzlinger, M, "Charge Transfer in Charge-Coupled Devices", J Appl Phys, 1971, 42, pp 3586-3594.

22. Carnes, J E and Kosonocky, W F, "Fast-Interface-State Losses in Charge-Coupled Devices", Appl Phys Lett, 1972, 20, pp 261-263.

23. Tompsett, M F, "The Quantitative Effects of Interface States on the Performance of Charge-Coupled Devices", IEEE Trans, Electron Devices, 1973, ED-20, pp 45-55.

24. Mohsen, A M, McGill, T C, Daimon, Y and Mead, C A, "The Influence of Interface States on Incomplete Charge Transfer in Overlapping Gate Charge-Coupled Devices", IEEE J Solid-State Circuits, 1973, SC-8, pp 125-138.

25. Vanstone, G F, Roberts, J B G and Long, A F, "The Measurement of the Charge Residual for CCD Transfer using Frequency and Impulse Responses", Solid-State Electronics, 1974, 17, pp 889-895.

26. Miller, G L, and Radeka, V, "Analogue Multiplication with Field-Effect Transistors", Proc Naval Postgraduate School Conf on Instrumentation Techniques in Nuclear Pulse Analysers, Monterey, California, 1963, pp 104-111.

27. Radeka, V, "Fast Analogue Multipliers with Field-Effect Transistors", IEEE Trans Nuclear Science, 1964, pp 302-307.

28. Mohsen, A M and Morris, F J, "Measurements on depletion-mode field-effect transistors and buried-channel MOS capacitors for the characterization of bulk transfer charge-coupled devices", Solid-State Electron., 1975, 18, pp407-416.

29. Frohman-Bentchkowsky, D and Grove, A S, "Conductance of MOS Transistors in Saturation", IEEE Trans Elect. Dev, 1969, ED-16, pp 108-113.

30. Frohman-Bentchkowsky, D and Vadasz, L, "Computer-Aided Design and Characterisation of Digital MOS Integrated Circuits", J Solid-State Circuits, 1969, SC-4, pp 57-64.

31. Armstrong, W E, "Charging efficiency of depletion load devices", IEEE J Solid State Circ., 1977, pp 313-316.

32. Tsividis, Y P, "Design considerations in single-channel MOS analogue integrated circuits - a tutorial", IEEE J Solid State Circ., 1978, SC-13 pp 383-391.

33. Shockley, W and Read, W T, "Statistics of the Recombinations of Holes and Electrons", Phys Rev, 1952, 87, pp 835-843.

34. Sangster, F L J and Teer, K, "Bucket Brigade Electronics - New Possibilities for Time Axis Conversion and Scanning", IEEE J Solid-State Circuits, 1969, SC-4, pp 131-136.

35. Browne, V A and Perkins, K D, "Buried Channel CCD's with Sub-micron Electrode Spacing", CCD 74, Dig of Tech Papers, Edinburgh, 1974, pp 100-105.

36. Walden, R H, Krambeck, R H, Strain, R J, McKenna, J, Schryer, N L and Smith, G E, "The Buried Channel Charge Coupled Device," Bell Syst Tech J, 1972, 51, pp 1635-1640.

37. Browne, V A, "Performance limitations of two phase CCD's", Agard Conference Proceedings No 230, 1978, p 2.7.

38. Thornber, K K, "Theory of noise in charge transfer devices", Bell Syst. Tech. J., 1974, 53, pp 1211-1262.

39. Carnes, J, Kosonocky, W F, "Noise Sources in Charge-Coupled Devices", RCA Rev, 1972, 33, pp 327-343.

40. ibid, "Measurements of Noise in Charge-Coupled Devices", ibid, 1973, 34, pp 553-565.

41. Klaassen, F M and Prins, J, "Thermal Noise of MOS Transistors", Philips Res Rep, 1967, 22, pp 505-514.

42. Chan, C H and Chamberlain, S G, "A CCD Serial to Parallel Shift Register", IEEE J Solid-State Circuits, 1973, SC-8, pp 388-391.

43. MacLennan, D J, Mavor, J, Vanstone, G F and Windle, D J, "Novel Tapping Technique for Charge-Coupled Devices", Electron Lett, 1973, 9, pp 610-611.

44. Hill, J R, McCaughan, D V, Keen, J M and White J C, "Charge Splitting CCD Tapped Delay Lines ", Electron. Lett., 1979, 15, pp 204-206.

45. Lampe, D R, White, M H, Mims, J H, Webb, W R and Gilmour, G A, "CCDs for Discrete Analogue Signal Processing (DASP)", presented at IEEE Int Convention and Exposition, 1974, Session 9.

46. Vanstone, G F, Harp, J G, MacLennan, D J and Mavor, J, "Analogue Correlators using Charge-Coupled Devices", Proc CCD App Conf, San Diego, 1975, pp 229-235.

47. Collins, D R, Bailey, W H, Gosney, W M and Buss, D D, "Charge-Coupled Device Analogue Matched Filters," Electron. Lett., 1972: 8, pp 328-329.

48. Tiemann, J J, Baertsch, R D, and Engeler, W E, "A Surface Charge Correlator for Signal Processing", CCD 73, Dig of Tech Papers, San Diego, 1973, pp 103-109.

49. Tiemann, J J, Engeler, W E and Baertsch, R D, "A Surface Charge Correlator", IEEE J Solid-State Circuits, 1974, SC-9, pp 403-410.

50. Lampe, D R, White, M H, Mims, J H and Fagan, J L, "An Electrically Programmable LSI Transversal Filter for Discrete Analogue Signal Processing", CCD 73, Dig of Tech Papers, San Diego, 1973, pp 111-125.

51. Lampe, D R, White, M H, Fagan, J L and Mims, J H, "An Electrically Programmable LSI Analog Transversal Filter", IEEE ISSCC, Dig of Tech Papers, Philadelphia, 1974, pp 156-157.

52. MacLennan, D J, Mavor, J and Vanstone, G F, "Technique for Realising Transversal Filters using Charge-Coupled Devices", Proc IEE, 1975, 122, pp 615-619.

53. White, M H, Lampe, D R, "Charge-Coupled Device Analog Signal Processing", Proc Int Conf on the Application of CCD's, San Diego, October 1975, pp 189-197.

54. Buss, D D, Collins, D R, Bailey, W H and Reeves, C R, "Transversal Filtering using Charge-Transfer Devices", IEEE J Solid-State Circuits, 1973, SC-8, pp 138-146.

55. Bosshart, P, 1976, "An integrated analogue correlator using charge-coupled devices", IEEE ISSCC 76, pp 198-199.

56. Mavor, J, Jack, M A, Saxton, D, and Grant, P M, "Design and performance of a programmable real-time CCD recirculating delay-line correlator", Electronic Circ. and Sys. 1977, 1, pp 137-144.

57. Haque, Y A and Copeland, M A, "Design and characterisation of a real-time correlator", IEEE J SSC, SC-12, 6, 1977, pp 642-649.

58. Evans, N E and Gamble, H S, "Intracell CCD programmable correlator", Elec. Lett., 1977, 13, 3, pp 69-71.

59. Denyer, P B, Arthur, J W and Mavor, J, "Monolithic, programmable analog CCD transversal filter", Elec. Lett., 1977, 13, pp 373-374.

60. Puckette, C M, Butler, W J and Smith, D A, "Bucket-brigade transveral filters", IEEE Trans. Commun., 1974, 22, pp 926-934.

61. Denyer, P B and Mavor, J, "Miniature programmable transversal filter using CCD/MOS technology", Proc. IEEE, 1979, 67, pp 42-50.

62. Gooding, J N, Curtis, T E, Pritchard, W D and Rehman, M A, "Programmable transversal filter using CCD components", Proc. Int. Conf. on App. of CCD, San Diego, 1978, pp 3B-23 to 3B-30.

63. Pritchard, W D and Gooding, J N, "Design and application of a cascadable binary weighted analogue correlator", Proc. Fifth Int. Conf. on CCD, Edinburgh, 1979, pp 241-246.

64. Weckler, G P and Walby, M D, "Programmable transversal filters: design tradeoffs", ibid., pp 211-221.

65. Denyer, P B and Mavor, J, "256-point programmable transversal filter", ibid., pp 253.

66. Sakane, T, Iida, T, Tanaka, F, Satoh, C and Suzuki, Y, "Digitally controlled and electrically programmable CCD transversal filter LSI", ibid., pp 222-229.

67. Chiang, A M, Burke, B E, Smythe, D L, Silversmith, D J and Mountain, R W, "A high speed CCD digitally programmable transversal filter", ibid., pp 230-236.

68. White, J C, Keen, J M, Homer, M F, McCaughan, D V and Hill, J R, "A fast 32 point analogue correlator", ibid., pp 237-240.

69. Tower, J R, Gandolfo, D A, Elliot, L D and McCarthy, B M, "A 512-stage analog-binary programmable transversal filter", ibid., pp 247-252.

70. Pelgrom, M J M, Wallinga, H and Holleman, J, "The electrically programmable splitted-electrode CCD transversal filter (EPSEF)", ibid., pp 254-260.

71. Denyer, P B and Mavor, J, "Monolithic 256-point programmable transversal filter", Elec. Lett., 1979, 15, pp 710-712.

72. Rabiner, L R and Gold, B, "Theory and Application of Digital Signal Processing", Prentice-Hall, New Jersey, 1975: Chapter 3.

73. McClellan, J H, Parks, T W and Rabiner, L R, "A computer program for designing optimum FIR linear phase digital filters", IEEE Trans. Audio Electroacoust., 1973, AU-21, pp 506-526.

74. Gold, B and Rader, C M, "Digital Processing of Signals", McGraw-Hill, 1969.

75. MacLennan, D J, "Design and Operation of a Programmable Charge-Coupled Device Transversal Filter", Thesis, University of Edinburgh, 1976. Chapter 6.

76. Fry, P W, "A MOST Integrated Differential Amplifier", IEEE J Solid-State Circuits, 1969, SC-6, pp 166-168.

77. Weste, N and Mavor J, "MOST amplifiers for performing peripheral integrated circuit functions", Electronic Circ. and Sys., 1977, 1, pp 165-172.

78. Tsividis, Y P and Gray P R, "An integrated NMOS operational amplifier with internal compensation", IEEE J Solid State Circuits, 1976, SC-11, pp 748-753.

79. Senderowicz, D, Hodges, D A and Gray, P R, "High performance NMOS operational amplifier", IEEE J Solid State Circuits, 1978, SC-13, pp 760-766.

80. Denyer, P B and Mavor,J, "Novel MOS differential amplifier for sampled-data applications", Electron. Lett., 1978, 14, pp 1-2.

81. Howes, M J and Morgan, D V, "Charge Coupled Devices and Systems," Wiley, 1979, Chapter 3.

82. Buss, D D, Bailey, W H and Collins, D R, "Analysis and applications of analog CCD circuits", Proc. 1973 IEEE Int. Symposium on Circ. Theory, pp 3-7.

83. Wolfson Microelectronics Liaison Unit, Edinburgh, "Development of CCD's for sonar signal processing - First progress report", MOD Contract N/CP11671/75/DC28GJR463, October 1976.

84. Wolfson Microelectronics Liaison Unit, Edinburgh, "Development of CCD's for sonar signal processing - Final report", MOD Contract N/CP11671/75/DC28GJR463, January 1978.

85. Wolfson Microelectronics Institute, Edinburgh, "Development of CCD's for sonar signal processing - First progress report", MOD Contract NCW/281/1128, February 1979.

86. Wolfson Microelectronics Institute, Edinburgh, "Development of CCD's for sonar signal processing - Final Report", MOD Contract NCW/281/1128, July, 1980.

87. Cowan, C F N, "Application of analogue charge coupled devices to Adaptive Signal Processing", Thesis, University of Edinburgh, 1980.

88. Cook, C E and Bernfeld, M, "Radar Signals", Academic Press, New York.

89. Cwik, A, Hobson, G S, Sitch, J E, Brewitt-Taylor, C R and Robson, P N, "Computer Simulation of Charge Coupled Devices", Proc. Fifth Int. Conf. on Charge Coupled Devices, Edinburgh, 1979, pp 412-416.

90. Bracewell, R, "The Fourier transform and its applications", McGraw-Hill, New York.

91. Tozer, R C and Hobson, G S, "Reduction of high-level non-linear smearing in CCDs", Electron. Lett., 1976, 12, pp 355-356.

92. Mavor, J, Davie, M C and Denyer, P B, "Techniques for increasing the effective charge transfer efficiency of tapped CCD registers", Electron. Lett., 1977, 13, pp 31-33.

93. Denyer, P B and Mavor, J, "Design of CCD delay lines with floating gate taps", SSED, 1977, 1, pp 121-129.

94. Fairchild CCD321 Application Note. 1977.

95. Wolfson Microelectronics Institute, WM2010 Application Note, Edinburgh, 1978.

96. Cowan, C F N, Arthur, J W, Mavor, J and Denyer, P B, "CCD based adaptive filters: realisation and analysis", IEEE Trans. ASSP, to be published.

97. Kapur, N, Mavor, J and Jack, M A, "Discrete cosine transform processor using a CCD programmable transversal filter", Electron. Lett., 1980, 16, pp 139-141.

98. Buss, D D, Bailey, W H, Holmes, J D, and Hite, L R, "Charge Transfer Devices; A new semiconductor technology applied to communication systems", Unpublished work. Texas Instruments Inc., Dallas, Texas.

APPENDIX I : PROCESS DATA

I.1 METAL GATE PROCESS

The metal gate process used for prototype designs, features a thin-field, two-metal technology, offering sub-micron intermetal gaps. These gaps are formed by the shadow etch technique illustrated in Figure I.1. Here the second metallisation is deposited over the first metallisation resist; because the first metallisation is slightly over-etched, the resist overhang ensures a uniform interelectrode gap.

A single depletion threshold, of -1v, is available, with p^+ channel-stopping necessary around all transistors, because of the uniform thin oxide. The second metallisation is also used as a field-plate over the entire active circuit area, except CCD, to block parasitic transistor action within the channel stop borders.

A summary of the mask steps and processing parameters is given in Table I.1.

The minimum feature size is generally 6 μm, although features may often need to be more widely separated for protection against electrical breakdown.

I.2 DOUBLE POLYSILICON GATE PROCESS

The polysilicon gate process offers two overlapping levels of polysilicon, which are separated only by a surface oxidation layer on the first level of polysilicon, see Figure I.2. The interelectrode gap is thus small ($\simeq$ .1μm), and electrically sealed. Self-aligned source and drain regions are a feature of this process, along with

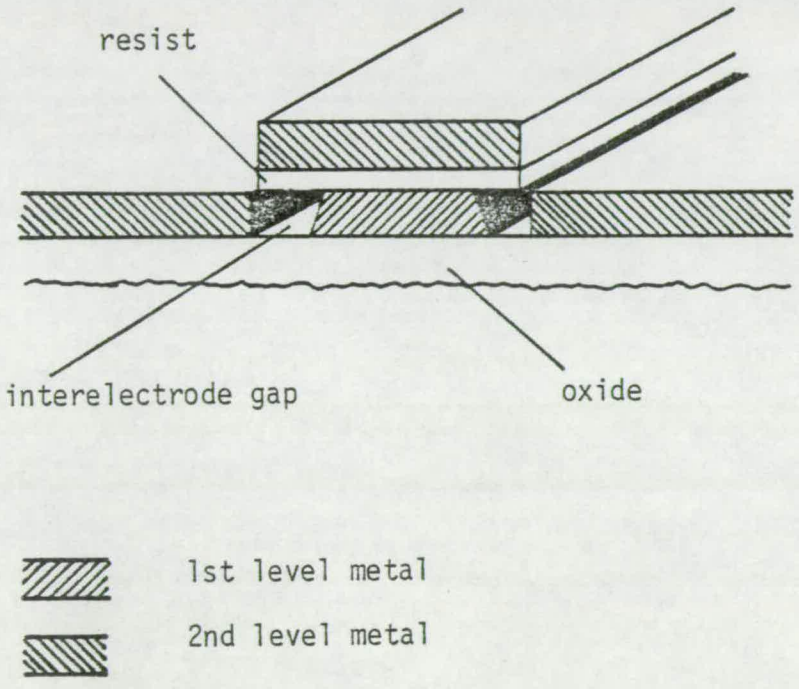


Figure I.1: Shadow-etch technique for achieving sub-micron interelectrode gaps.

MASKS: n^+ diffusion

p^+ channel-stop

Aluminium 1

Aluminium 2

Contacts

Pyro

PROCESS PARAMETERS:

$V_{T_o} = -1v$

$\beta_o = 15\ \mu A/v^2$

$\gamma = 0.33$

$V_{T_{p+}} = 25v$

$C_{ox} = 2.4 \times 10^{-4}\ pF/\mu^2$

$C_{n+} = 2.0 \times 10^{-5}\ pF/\mu^2$ (Area)

$7.0 \times 10^{-5}\ pF/\mu^2$ (Periphery)

Table I.1: Summary of SET Process

a self-aligned thick-field oxide and implant, which remove the channel-stop layout problem. The minimum feature size is 5μm, with 2μm polysilicon overlap. These dimensions combine to set the minimum (4-electrode) CCD cell pitch at 28μm.

The second polysilicon level receives a self-aligned enhancement implant for 2-phase CCD operation, and both polysilicon levels may receive an optional depletion implant. Thus a range of four transistor thresholds emerges.

A summary of the mask steps and processing parameters is given in Table I.2.

The process is intended to offer 15V operation for clocks and supplies, with a -5V substrate bias.

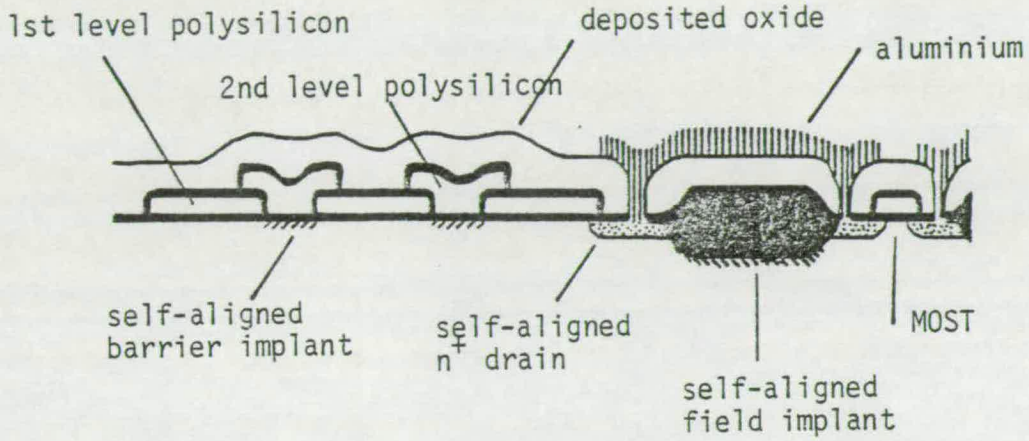


Figure I.2: Schematic cross-section of coplanar polysilicon gate process

MASKS: Active Area

Depletion Area

Polysilicon Level 1

Polysilicon Level 2

Contacts

Aluminium

Passivation

TRANSISTOR PARAMETERS:

	Level 1 Enhancement	Level 1 Depletion	Level 2 Enhancement	Level 2 Depletion
V_{T-5}(V)	1.25	-9	3.5	-5
β_o(A/V)	25	17	17	14
γ	0.3	0.5	0.7	1.0
θ	0.05	-0.04	-	-

PROCESS PARAMETERS:

$C_{OX} \simeq 3.3 \times 10^{-4}$ pF/μ^2

$C_{p-p} \simeq 1.8 \times 10^{-4}$ pF/μ^2

$C_{A-P} \simeq 3.3 \times 10^{-5}$ pF/μ^2

$C_{n+} \simeq 3.0 \times 10^{-5}$ pF/μ^2 (Area)

3.0×10^{-5} pF/μ^2 (Periphery)

Table I.2: Summary of double polysilicon gate process

APPENDIX II : LINEAR MOST STAGES

II .1 A SMALL SIGNAL MODEL

A practical small-signal model of the MOS transistor is shown in Figure II.1. Changes in channel current are stimulated through the gate, or forward, transconductance, g_m, and through the channel conductance, g_{ds}. Evaluation of these parameters for the first order transistor model gives;

- In the triode region

$$g_m = \beta V_{DS} \qquad \text{(II .1)}$$

$$g_{ds} = \beta(V_E - V_{DS}) \qquad \text{(II .2)}$$

- In saturation

$$g_m = \beta V_E \qquad \text{(II .3)}$$

$$g_{ds} = 0 \qquad \text{(II .4)}$$

These equations may be modified where necessary to include relevant second order terms.

Of the capacitive components, those between the source/drain, and the substrate, represent depletion capacitances associated with these reverse-biased diodes; they will normally be lumped with other

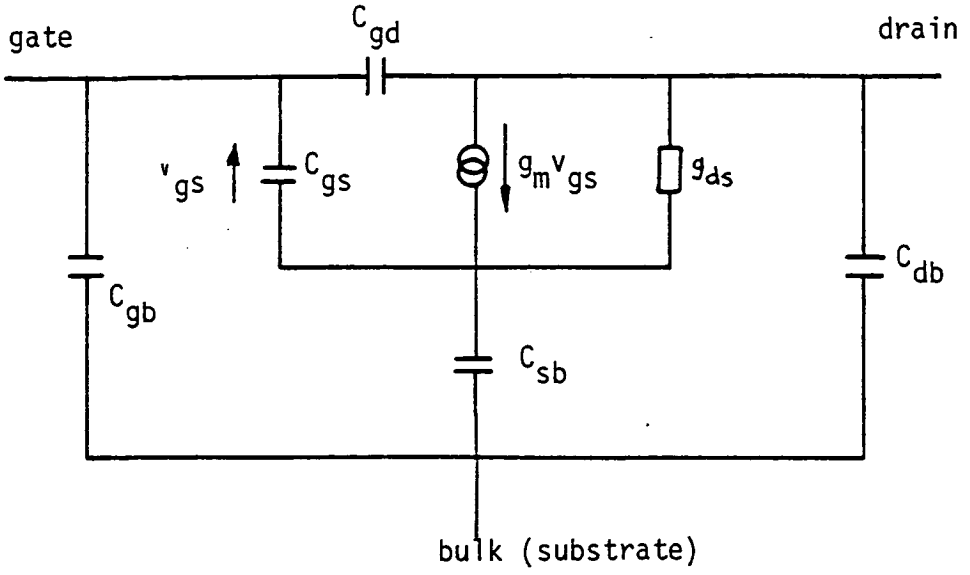


Figure II.1: Small signal MOST model

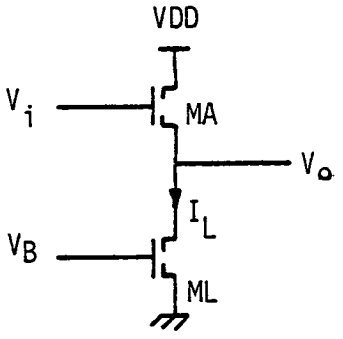


Figure II.2: MOST source follower with constant current load

elements connected to these nodes. The various gate components are somewhat complex, in that the capacitive values are not constant, but vary with the bias conditions. The values are approximately constant for devices biased into saturation, and are given by[6];

$$C_{gd} = C_{gd_e} \qquad \text{(II .5)}$$

$$C_{gs} = C_{gs_e} + \frac{2}{3} C_i \qquad \text{(II .6)}$$

$$C_{gb} \simeq 0.1\ C_i, \qquad \text{(II .7)}$$

where e denotes the external, or overlap, component, and i denotes the internal, or intrinsic, component.

II .2 THE SOURCE FOLLOWER

The most common linear application of the MOS transistor is in the voltage buffer configuration shown in Figure II.2. The stage has a high input impedance and a low output impedance, and is used to buffer signals from high impedance sources (often purely capacitive).

The transistor ML, operating in *saturation*, forms a constant current load, of value I_L. In order to satisfy this current demand the active transistor MA, again in saturation, must maintain a constant effective gate voltage. Taking the first order model of (3.15), it is clear that the source voltage must track the gate, so achieving a small-signal through-gain of unity. Note that a d.c. offset,

$$V_I - V_0 = V_{T_A} + \sqrt{\frac{2I_L}{\beta_A}} \qquad \text{(II .8)}$$

is maintained.

Considering the relevant second-order effects; channel-length modulation in both saturated transistors, and body effect in the active transistor, all reduce the gain of the stage, as well as modifying slightly the anticipated offset. Channel-length modulation effects are normally reduced by using a lengthened channel geometry, but there is no reduction of the body effect, which reduces the through gain to[32];

$$A_{V_0} = \frac{\partial V_0}{\partial V_I} = 1 - \frac{\partial V_{T_A}}{\partial V_I}$$

$$\simeq \frac{1}{1 + \frac{\gamma}{2\sqrt{V_{SB}}}} \; . \qquad \text{(II .9)}$$

For typical processes A_{V_0} lies in the range 0.9 to 0.95.

Because the active transistor source tends to follow the gate, any C_{gs} component becomes partially cancelled. Thus the input capacitance of this stage is low;

$$C_{in} = C_{gd_e} + C_{gb} + C_{gs}(1 - A_{V_0}) \qquad \text{(II .10)}$$

The output conductance of the source follower stage with constant-current load is given by;

$$g_o = g_{m_A} + g_{ds_A} + g_{ds_L} \qquad \text{(II .11)}$$

where, for transistors operating in saturation, $g_{m_A} >> g_{ds_{A,L}}$ and thus;

$$g_o \simeq g_{m_A} \simeq \sqrt{2\beta_A I_L} \qquad \text{(II .12)}$$

II .2 THE MOST INVERTER

As in bipolar stages, the MOST inverter forms a useful gain block. High values of resistance are not feasible in LSIMOS however, so that MOST's must also be used as load devices. This leads to drift-sensitive stages whose gain and bias characteristics are not easily defined, or controlled. For these reasons, the inverter is normally only used to provide gain within a controlled feedback loop, often as part of an operational amplifier.[32]

Three different invertor circuits are shown in Figure II.3. The first two are enhancement-and depletion-load versions of the same simple configuration, consisting of an active transistor MA, and a load device ML. The third connection incorporates two improvements; cascode-connected transistor MC blocks output voltage changes from the drain of the active transistor at node A, thus cancelling the Miller effect which becomes serious at high gains, and current-source MI boosts the transdconductance of MA to increase the stage gain.

The gain of all three stages is given by the product of the active transistor transconductance with the effective resistance of the load. Thus;

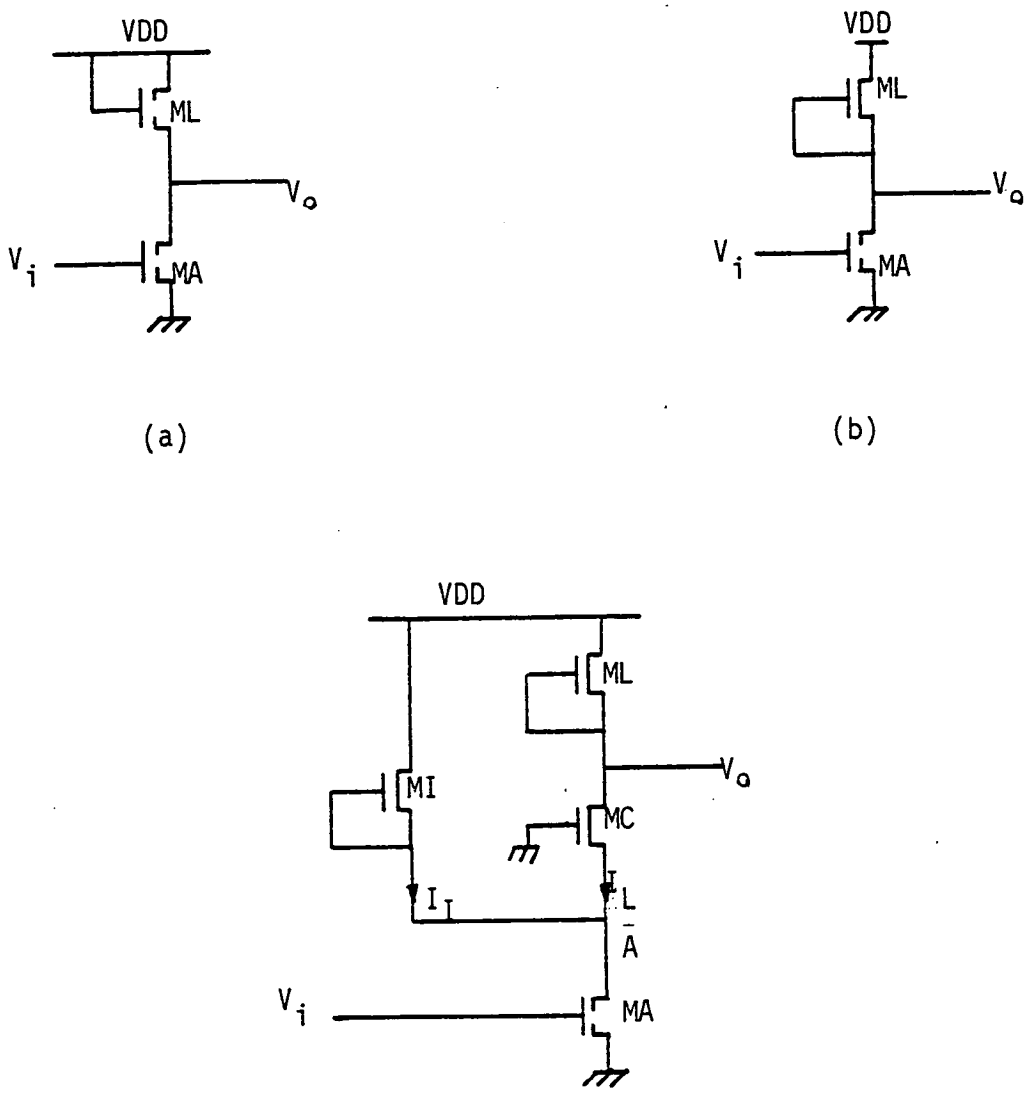


Figure II.3: MOST Inverter configurations

(a) enhancement load
(b) depletion load
(c) cascode-connected depletion load

$$A_{V_0} = g_{m_A} r_L \qquad (II.13)$$

Assuming that long channel geometries are chosen to minimise the effects of channel-length modulation, the output impedance of these stages is given approximately by the effective load resistance r_L. Finally the input capacitance is given by;

$$C_{in} = C_{gs} + C_{gb} + C_{gd}(-A_{V_D}), \qquad (II.14)$$

where A_{V_D} is the small-signal voltage gain at the active transistor drain.

Table II.1 summarises these properties for each invertor configuration to first order , and gives typical values of three parameters, assuming Aspect Ratios of 10 and 0.1 are used for the active and load devices respectively.

Clearly, for the extra layout area and design effort involved, the cascode-connected depletion-load inverter offers a considerable reduction in input capacitance, coupled with an increase in gain. The simple depletion load inverter does offer an increase in gain over the enhancement load, but only at the sacrifice of considerably increased input capacitance.

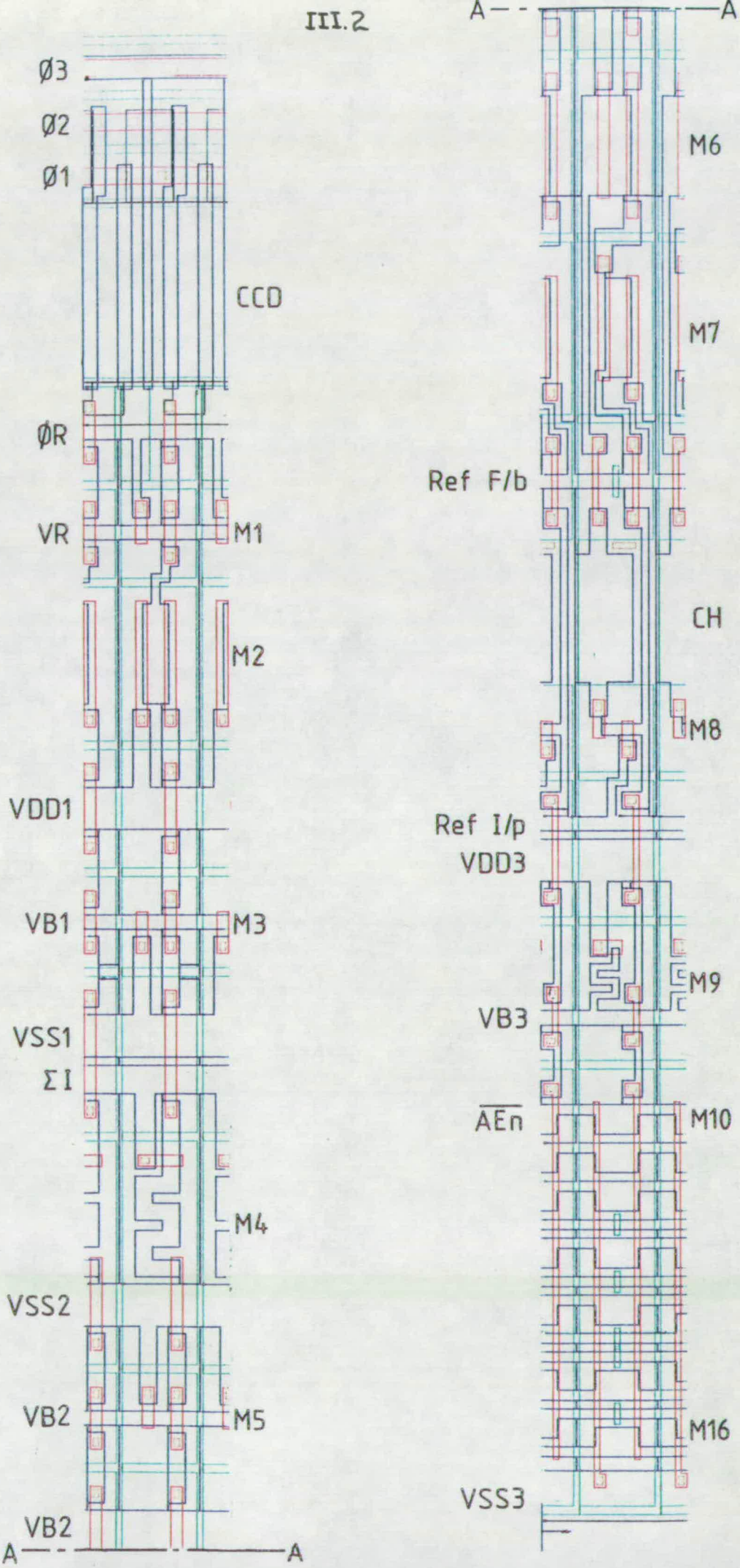
Stage	C_{IN}	C_{IN} typ.	A_{VO}	A_{VO} typ.	r_o	r_o typ.
Enhancement load	$C_{gs} + C_{gb}(-A_{V_0})$	.3pF	$\frac{\beta_A}{\beta_L}$	-10	$\frac{1}{\beta_L V_{E_L}}$	100K
Depletion Load	$C_{gs} + C_{gb} + C_{gd}(-A_{V_0})$	2pF	$\frac{2\sqrt{V_0}}{\gamma}\sqrt{\frac{\beta_A}{\beta_L}}$	-100	$\frac{2\sqrt{V_0}}{\gamma\beta_L V_{E_L}}$	1M
Cascode-connected Depletion Load	$C_{gs} + C_{gb} + C_{gd}(-A_{V_D})$	.1pF	$\frac{2\sqrt{V_0}}{\gamma}\sqrt{\frac{\beta_A}{\beta_L}}\frac{I_I}{I_L}$	-500	$\frac{2\sqrt{V_0}}{\gamma\beta_L V_{E_L}}$	1M

Table II.1: Summary of Inverter Performance Expressions (approximate).

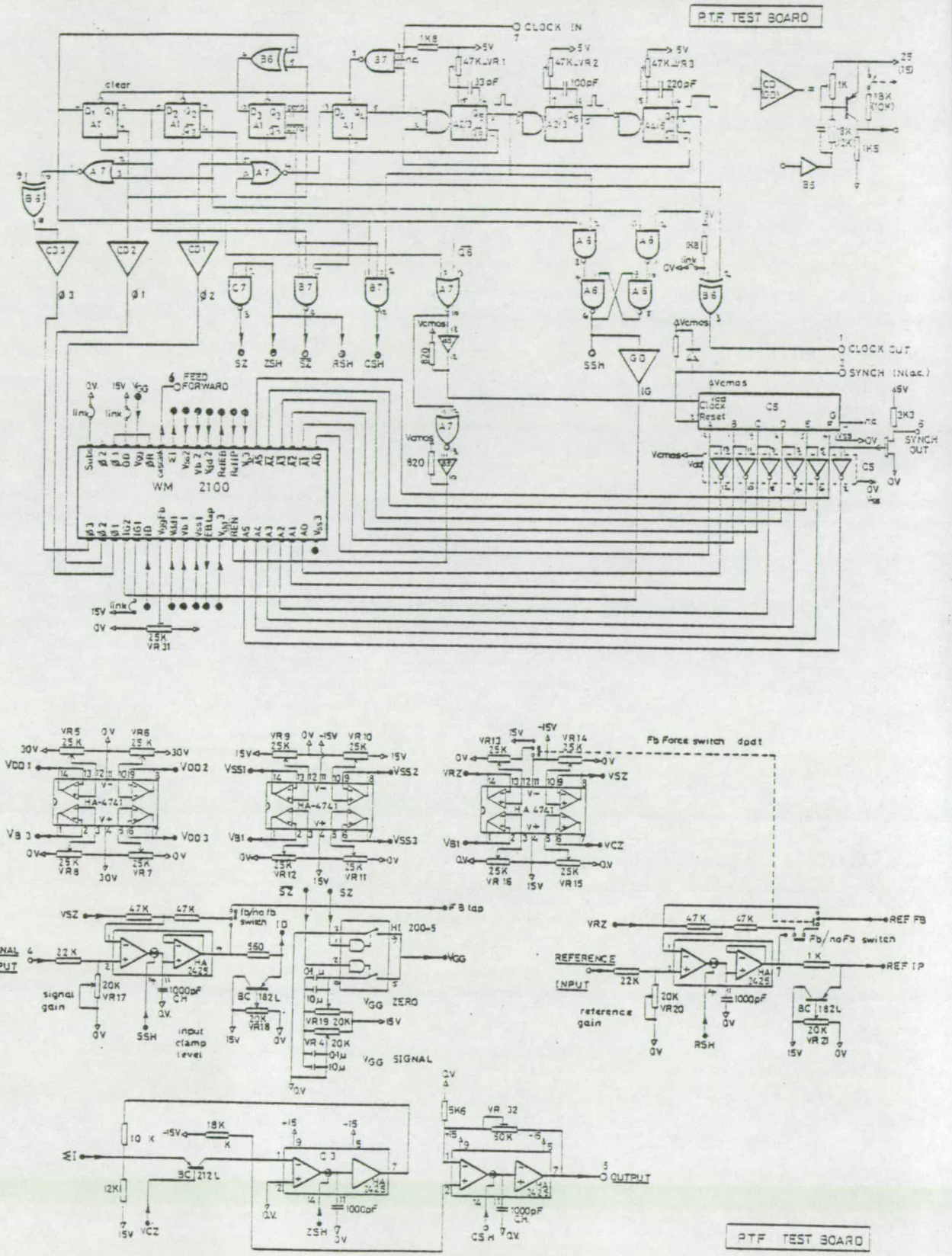
APPENDIX III : PTF I CELL LAYOUT

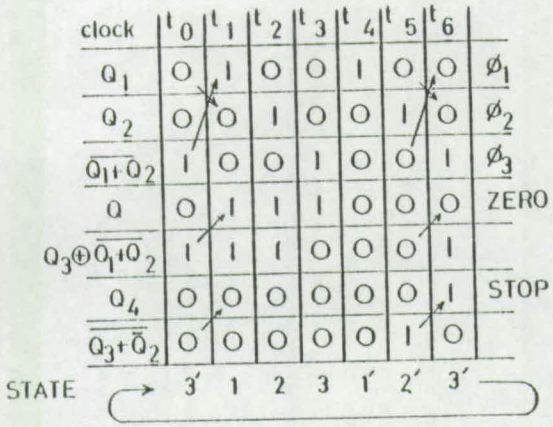
Scale : 1" = 100μm

Key :		
	p^+ channel-stop	Green
	n^+ diffusion	Red
	Aluminium I	Black
	Aluminium II	Blue
	Contact Windows	Brown


Ø3
Ø2
Ø1
CCD
ØR
VR
M1
M2
VDD1
VB1
M3
VSS1
ΣI
M4
VSS2
VB2
M5
VB2
A
A
A
A
M6
M7
Ref F/b
CH
M8
Ref I/p
VDD3
M9
VB3
AEn
M10
M16
VSS3

APPENDIX IV : CIRCUIT AND TIMING DIAGRAMS FOR PTF I PROTOTYPE BOARD

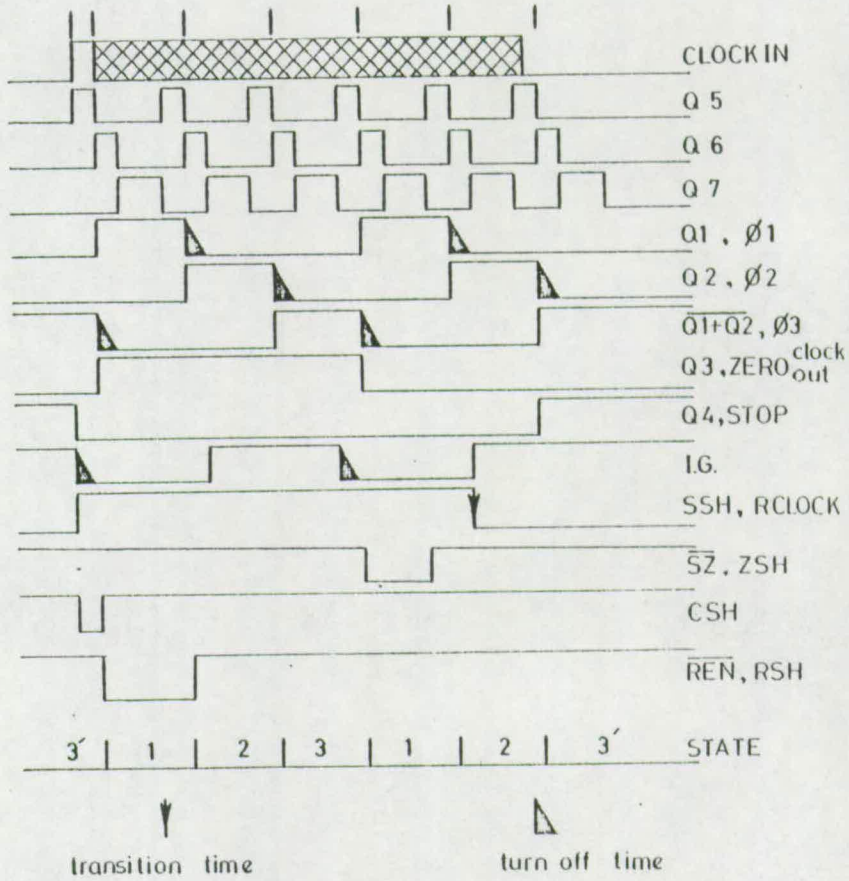

P.T.F. TEST BOARD
CLOCK IN
clear
CLOCK OUT
SYNCH (Nlac.)
SYNCH OUT
FEED FORWARD
WM 2100
SZ
ZSH
SSH
RSH
CSH
VR31
VR8
VR7
VR12
VR11
VR16
VR15
VR17
VR18
VR19
VR4
VR20
VR21
VR32
SIGNAL INPUT
signal gain
input clamp level
VGG ZERO
VGG SIGNAL
REFERENCE INPUT
reference gain
Fb Force switch dodt
Fb/noFb switch
OUTPUT
P.T.F. TEST BOARD



STATE DIAGRAM

Each state is divided into three consecutive times, Q_6, Q_7 and Q_5. Q_5 defines the period within which the input gate must be turned off before a change of state. Q_6 defines the state transition time in which one or other of the clocks ϕ_1, ϕ_2 or ϕ_3 is turning off. Q_7 defines the time during which the clocks and input gate may be taken as being stable.

PTF TEST BOARD STATE DIAGRAM

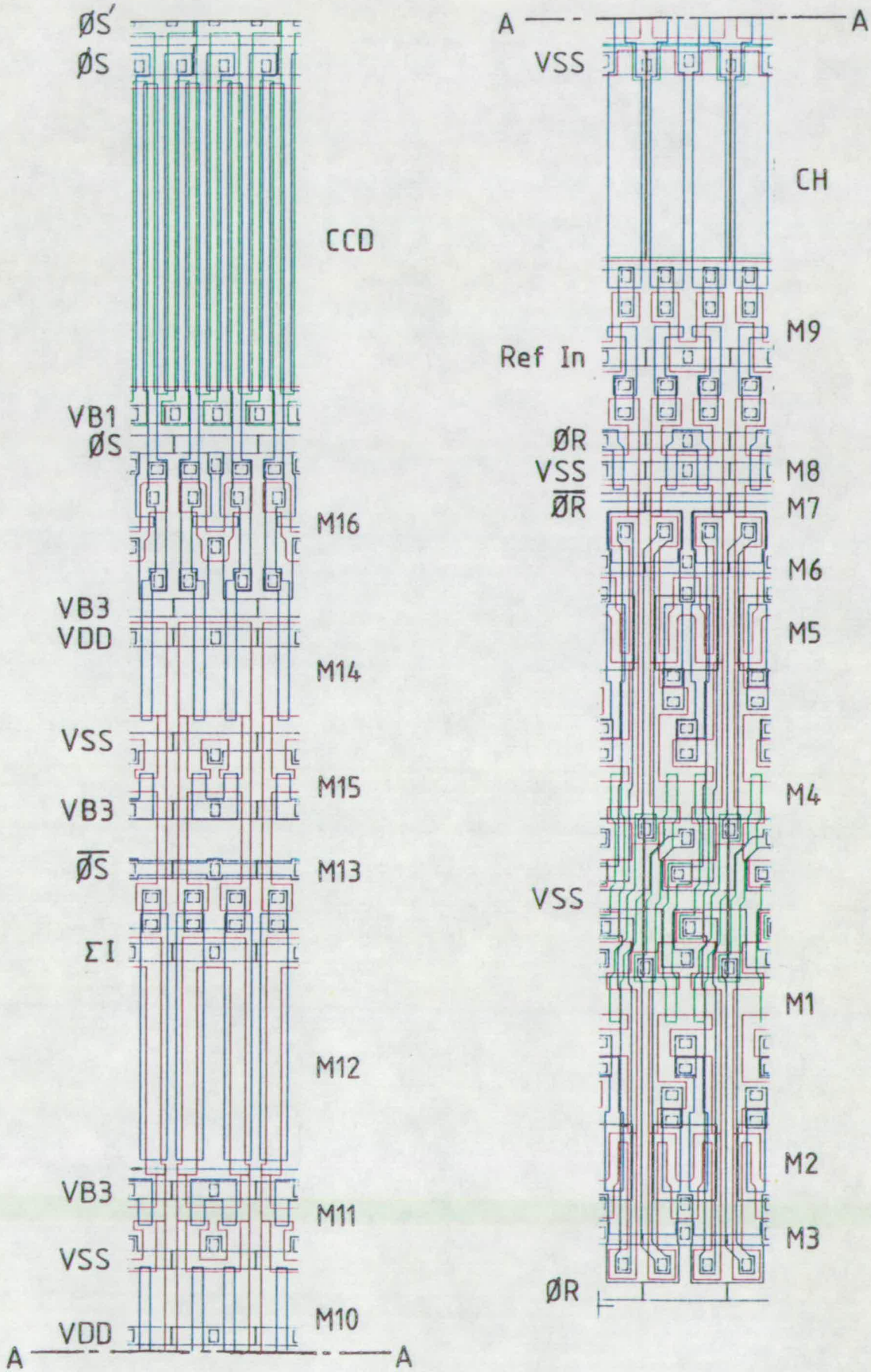


PTF TEST BOARD TIMING DIAGRAM

APPENDIX V : PTF II CELL LAYOUT

Scale : 1" = 100μm

Key :		
	Active area	Red
	Depletion region	Turquoise
	Polysilicon I	Blue
	Polysilicon II	Green
	Contact windows	Black
	Aluminium	Black


ØS'
ØS
CCD
VB1
ØS
M16
VB3
VDD
M14
VSS
M15
VB3
ØS
M13
ΣI
M12
VB3
M11
VSS
VDD
M10
A
A
A
A
VSS
CH
M9
Ref In
ØR
VSS
M8
ØR
M7
M6
M5
M4
VSS
M1
M2
M3
ØR

APPENDIX VI : LIST OF AUTHOR'S PUBLICATIONS

*Those items marked * are bound at the end of this Thesis.*

1. "Techniques for increasing the effective charge-transfer efficiency of tapped CCD registers", Electronics Letters, 1977, 13, pp. 31-33.

* 2. "Design of CCD delay lines with floating gate taps", Solid State and Electron Devices, 1977, 1, pp 121-129.

3. "Analogue CCD correlator using monolithic MOST multipliers", Electronics Letters, 1977, 13, pp 373-374.

4. "The design and development of CCD programmable transversal filters and correlators", AGARD conference on impact of CCD and SAW devices on signal processing and imagery in advanced systems, Ottawa, 1977, pp 2-8.

* 5. "Design and development of CCD programmable transveral filters", Electronic Circuits and Systems, 1978, 2, pp 1-8.

* 6. "Novel MOS differential amplifier for sampled data applications", Electronics Letters, 1978, 14, pp 1-2.

* 7. "Monolithic programmable analogue CCD transversal filter", Electronics Letters, 1977, pp 739-741.

8. "The dependence of CCD dark current upon power dissipation", Microelectronics and Reliability, 1978, 17, pp 403-404.

9. "Monolithic 64-point programmable analogue CCD/MOS transversal filter", European Solid State Circuits Conference, 1978.

10. "A programmable CCD transversal filter; design and application", Fourth International Conference on Charge Coupled Devices, San Diego, 1978.

11. "A CCD based programmable transversal filter", IEE Colloqium on Electronic Filters, London, 1978.

12. "Design and development of a CCD programmable transversal filter", IEE Colloqium on Sonar, London, 1978.

* 13. "Miniature programmable transversal filters using CCD/MOS technology", Proceedings of the IEE, 1979, 67, pp 42-50.

* 14. "A monolithic CCD programmable transversal filter for analogue signal processing", The Radio and Electronic Engineering, 1980, 50, pp 213-225.

15. "CCD based adaptive filters: realisation and analysis", IEEE Trans. Acoustics Speech and Signal Processing, to be published.

16. "An evaluation of analogue and digital adaptive filter realisations", Case Studies in Advanced Signal Processing, Peebles, 1979.

17. "Large time bandwidth product CCD correlators for sonar", Case Studies in Advanced Signal Processing, Peebles, 1979.

18. "CCD based analogue adaptive processing", Fifth International Conference on Charge Coupled Devices, Edinburgh 1979.

19. "256-point programmable transversal filter", Fifth International Conference on Charge Coupled Devices, Edinburgh, 1979.

* 20. "Monolithic 256-point programmable transversal filter", Electronics Letters, 1979, 15, pp 710-712.

TECHNIQUES FOR INCREASING THE EFFECTIVE CHARGE-TRANSFER EFFICIENCY OF TAPPED C.C.D. REGISTERS

Indexing terms: Charge-coupled-device circuits, Delay lines

A design technique for multitap c.c.d. delay lines is discussed in which the effective charge transfer efficiency is increased over its intrinsic process-dependent value. The technique involves locating tap amplifiers at every alternate bit, and operating the device at twice the normal clock rate. The advantages of the technique are discussed with reference to a 32-tap, *n*-channel c.c.d. delay line.

Introduction: Recent work has shown how an improvement in the effective transfer efficiency of c.c.d.s may be obtained by the introduction of cell redundancy and circuit complexity.[1, 2] Techniques reported here employ cell redundancy, but involve a minimum of peripheral circuit complexity; these are especially suitable for multitap delay lines, as well as single-output registers.

Consider a c.c.d. register operated in the conventional mode, as shown in Fig. 1*a*. The impulse-response sequence, allowing for transfer inefficiency, has been well studied and an adaption of the result obtained by Vanstone[3] is used here. The *r*th residual of the impulse response sequence at a non-destructive tap *n* can be shown to be

$$\frac{(n+r)!}{r!n!}\varepsilon^r(1-\varepsilon)^n \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (1)$$

where ε is the effective transfer inefficiency per cell and $r = 0$ indicates the main charge packet, $r = 1$ the first residual etc.

The effect of ε is to smear a single charge packet into following signal samples. For low $n\varepsilon$ products, this is limited to a predominant first residual contribution to the immediately following signal sample. This letter discusses two circuit techniques which may be used to reduce the effect of transfer inefficiency, and results are presented for a multitap c.c.d. delay line designed to employ these principles.

Description: The first scheme, shown in Fig. 1*b*, employs alternate input sampling, in which 'fat zeros' are interposed between the signal packets. By sampling only the signal packets at output taps, the contribution of the preceding signal sample is reduced from a first to second residual effect, the intervening 'zero' having absorbed the comparatively large first residual.

The second scheme, shown in Fig. 1*c*, provides self cancellation of the first residual loss, as well as a reduction in the contribution of preceding signal packets. Each input signal sample is injected in two successive charge packets, and the second charge packet of each pair is sampled at output taps. The reduction of the second packet by transfer inefficiency is compensated by the addition of the (ideally) identical loss of the leading packet during each transfer. The leading charge packet also reduces the residual effect of preceding signal samples.

As both schemes halve the effective data rate, it is necessary to double the clock frequency and the number of stages to achieve the same sampling criteria and time–bandwidth product; the halfrate signal charge packets may be suitably sampled at alternate stages. For convenience, the schemes are referred to as alternate zero alternate tap (a.z.a.t.) and double sample alternate tap (d.s.a.t.), respectively.

The impulse-response sequence of the a.z.a.t. scheme may be obtained from the conventional response (expr. 1), considering alternate terms only and allowing for $2n$ stages to tap n:

$$\frac{(2n+2r)!}{2r!2n!}\varepsilon_2^{2r}(1-\varepsilon_2)^{2n} \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (2)$$

where ε_2 is the value of ε at the new clock frequency $2f_s$ (ε_1 will be taken as the value of ε at f_s). That of the d.s.a.t. scheme is obtained by considering alternate terms of the conventional response to two impulses delayed by one clock period with respect to one another, again allowing for $2n$ stages:

$$\frac{(2n+2r)!}{2r!2n!}(1+2n\varepsilon_2/(2r+1)+\varepsilon_2)\varepsilon_2^{2r}(1-\varepsilon_2)^{2n} \quad . \quad (3)$$

A summary of these results is given in Table 1, where the techniques are compared in terms of a quality factor R, defined as the magnitude ratio of the first residual to the main charge packet.

The results for the d.s.a.t. scheme are identical to those achieved by the scheme proposed by Tozer and Hobson;[1] in fact, the principle is similar. However, addition of the main charge packet and its first residual is here accomplished automatically during each transfer, rather than by peripheral circuitry.

Limitations and comparisons: The benefit obtained from these techniques is reduced at high operating frequencies where doubling the clock frequency may degrade ε_2 significantly. Indeed, an upper limit on the operating frequency may be determined by the criterion

$$\varepsilon_2^2 < \varepsilon_1/2n \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (4)$$

Clearly, the schemes perform best at clock frequencies where ε is effectively constant; it is convenient to compare the results under these conditions.

Table 1 shows that, where $\varepsilon_2 \simeq \varepsilon_1$, both schemes offer an improvement over conventional operation. For practical $n\varepsilon$ values, the improvement in quality factor R is approximately $1/2n\varepsilon$. Comparison of the two proposed techniques shows that double sampling provides a marginal performance improvement on alternate zero operation. Reference to Fig. 1 also shows that implementation of the d.s.a.t. scheme is slightly simpler. It is thus concluded that double sampling is preferable to alternate zero operation.

Device considerations: The increased number of transfer cells appears initially to be a disadvantage. For multitap-register applications, however, the increase in available silicon area may be advantageous where posttap signal processing is required on chip. The distance between tap outputs is often

Table 1 COMPARISON OF THREE TECHNIQUES

	Impulse response (residual r) (tap n)	$r_1 : r_0$ performance ratio		
		General	Large n	$n\varepsilon = 0{\cdot}1$
Conventional	$\frac{(n+r)!}{r!n!}\varepsilon_1^r(1-\varepsilon_1)^n$	$\varepsilon_1(n+1)$	$n\varepsilon_1$	0·100
Alternate zero alternate tap	$\frac{(2n+2r)!}{2r!2n!}\varepsilon_2^{2r}(1-\varepsilon_2)^{2n}$	$\varepsilon_2^2(n+1)(2n+1)$	$2(n\varepsilon_2)^2$	0·020
Double sample alternate tap	$\frac{(2n+2r)!}{2r!2n!}(1+2n\varepsilon_2/(2r+1)+\varepsilon_2)\varepsilon_2^{2r}(1-\varepsilon_2)^{2n}$	$\varepsilon_2^2(n+1)(2n+1)\frac{(1+\varepsilon_2(n+1))}{(1+\varepsilon_2(2n+1))}$	$2(n\varepsilon_2)^2(1-n\varepsilon_2)$	0·018

Reprinted from ELECTRONICS LETTERS 6th January 1977 Vol. 13 No. 1 pp. 31–33

determined by transfer efficiency and process considerations and can be restrictive where identical signal-processing circuits are required at every tap. The schemes described here may be used to double the area available for peripheral signal processing circuitry, for a given c.c.d. cell length.

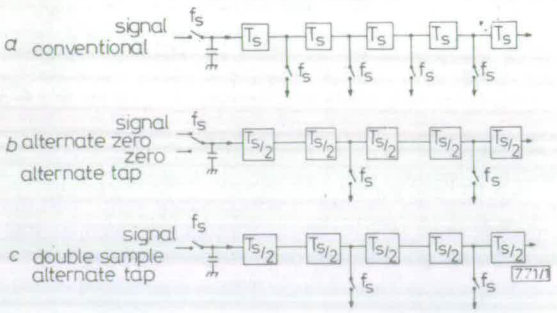


Fig. 1 *Three modes of c.c.d. operation*
f_s = sample frequency
T_s = sample period

Conversely, where the distance between taps is determined by posttap circuitry, the schemes allow gate lengths to be halved for a given circuit configuration, permitting higher-frequency operation.

Implementation of both schemes involves little peripheral complexity, the only additional circuit requirement being the generation of sample pulses at half the clock frequency. This practical simplicity makes the schemes very attractive where an improvement in efficiency or increase in available peripheral circuit area is desired without loss in device performance. Where the increase in clock frequency and device length is impractical, it is possible to implement the schemes and preserve the lower clock frequency by multiplexing two parallel registers.

The principle may be extended to include higher-order sampling and cell redundancy where greater improvements in efficiency or available peripheral circuit area are necessary.

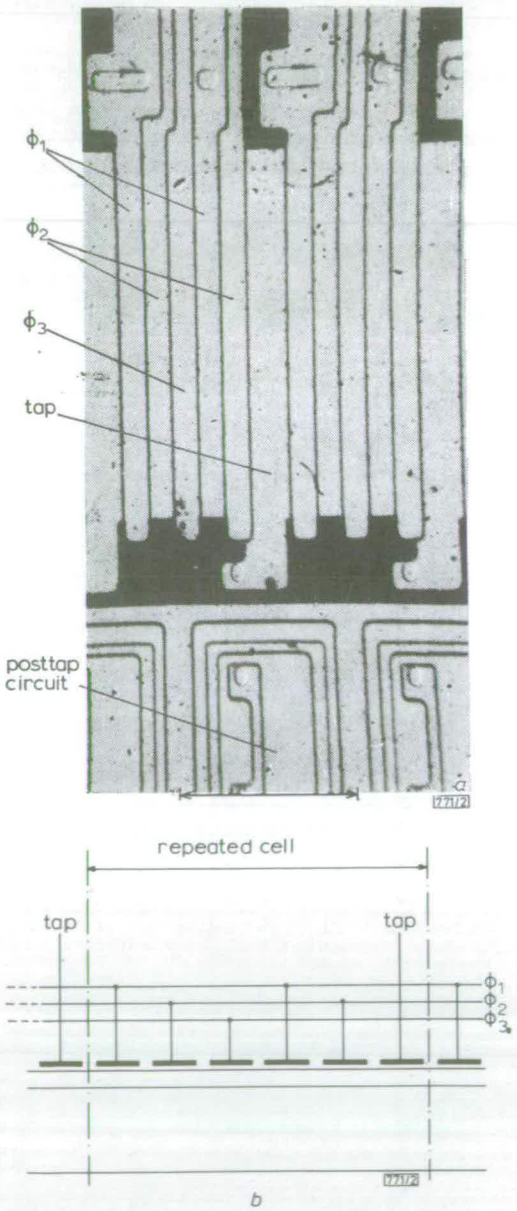


Fig. 2 *Experimental device*
b Photomicrograph of c.c.d.
a Schematic of cell structure

Experimental results: A 64-bit (32-tap) c.c.d. and its peripheral f.g.r.[4] tapping circuitry (Fig. 2) was fabricated with a 'shadow-etch' (s.e.t.) process.[5] The device has gate lengths of 5 μm, with 10 μm tap gates and tap sense amplifiers of 35 μm pitch.

The device was operated with a fill-and-spill input technique. Fig. 3*a* shows the sampled output at tap 13 in response to a pulse input (shorter in duration than the clock period) which was adjusted to give 90% of full well capacity. The quality factor R is estimated from the photograph to be 0·04. Fig. 3*b* shows the corresponding output at tap 26, employing the d.s.a.t. technique, and quite clearly a significant improvement in the quality factor has been achieved; in fact, the improvement is such that the effective first residual is difficult to measure. Theoretically, the improved quality factor is approximately 0·003. It is interesting to note that the main response in Fig. 3*b* is the sum of the main response and the first residual in Fig. 3*a*, as predicted by eqn. 3.

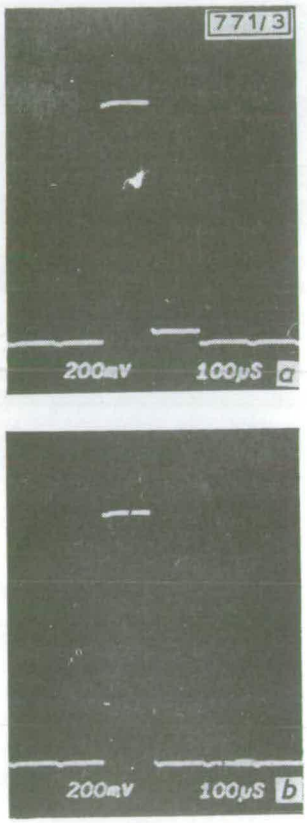


Fig. 3 *Impulse responses*
a Normal operation ($f_c = f_s = 10$ kHz)
b Double sampling (d.s.a.t.) operation ($f_c = 20$ kHz, $f_s = 10$ kHz)

Conclusions: Two techniques have been discussed for improving the effective charge transfer efficiency of multitapped c.c.d. delay lines. Both can be implemented with little increase in peripheral circuitry. The apparent disadvantages of the techniques, twice the clock frequency and device length, should be outweighed in the majority of applications requiring high-efficiency values. The increases in device area may help the layout of the tap amplifiers. The d.s.a.t. technique is marginally better than the a.z.a.t. approach and the efficacy of the former technique has been demonstrated for a 32-bit c.c.d. delay line.

Acknowledgments: The authors gratefully acknowledge the support of the UK SRC, DCVD MOD(PE) and the Wolfson Microelectronics Liaison Unit, University of Edinburgh, for device design facilities. The devices were designed by D. J. MacLennan and fabricated at The Plessey Co. Ltd.

J. MAVOR
M. C. DAVIE

Department of Electrical Engineering
School of Engineering Science
University of Edinburgh
Edinburgh EH9 3JL, Scotland

P. B. DENYER

Wolfson Microelectronics Liaison Unit
School of Engineering Science
Mayfield Road, Edinburgh EH9 3JL, Scotland

22nd November 1976

References

1 TOZER, R. C., and HOBSON, G. S.: 'Reduction of high-level nonlinear smearing in c.c.d.s', *Electron. Lett.*, 1976, **12**, pp. 355–356
2 COOPER, D. C., DARLINGTON, E. H., PETFORD, S. M., and ROBERTS, J. B. G.: 'Reducing the effect of charge-transfer inefficiency in a c.c.d. video integrator', *ibid.*, 1975, **11**, pp. 384–385
3 VANSTONE, G. F., ROBERTS, J. B. G., and LONG, A. E.: 'The measurement of the charge residual for c.c.d. transfer using impulse and frequency responses', *Solid-State Electron.*, 1974, **17**, pp. 889–895
4 MCLENNAN, D. J., MAVOR, J., VANSTONE, G. F., and WINDLE, D. J.: 'Novel tapping technique for charge-coupled devices', *Electron. Lett.*, 1973, **9**, pp. 610–611
5 PERKINS, K. D., and BROWNE, V. A.: 'Sub-micron gap metal gate technology for CCDs', *Microelectron.* 1975, **7**, (2), pp. 14–22

Design of c.c.d. delay lines with floating-gate taps

P.B. Denyer and J. Mavor

Indexing terms: *Charged-coupled device circuits, Delay lines*

Abstract: Multitapped c.c.d. analogue delay lines have been produced with the floating-gate, reset-sensing technique. Although the efficacy of the approach has been demonstrated, no comprehensive design procedure exists to enable systematic device design. Because the c.c.d. and its associated tapping circuitry is an active structure, the operational parameter relationships are extremely complex and dependent on many physical effects. Some of these individual processes have been previously associated with a particular operating parameter, but, usually, for a nontapped device configuration. This paper summarises the basic performance limiting processes of floating-gate tapped c.c.d. delay lines, and presents a quantitative basis for designs and also for further analytical studies. In particular, 3-phase surface-channel devices are considered, although the analyses may be extended to other c.c.d. formations. The equations presented are related to a simple design example based upon a specification achievable in practical devices.

1 Introduction

A useful configuration for an analogue mode charge-coupled device (c.c.d.) is the serial-in parallel-out mode, in which a number of sense amplifiers are used to detect the magnitude of sequentially injected signal-charge 'packets'. A usual constraint on these amplifiers is that the charge must be sensed nondestructively, without impeding its progress along the c.c.d. delay line. Additionally, the sensing technique should not add noise or distortion to the signal being detected. A necessary requirement for the sensing structure is that it can be integrated in c.c.d. technology and occupy a minimum of layout area. M.O.S. transistors are process compatible with the c.c.d. and are usually employed to form the tapping circuitry.

The floating-gate reset (f.g.r.) tapping structure[1] is now recognised as a useful sensing circuit for many c.c.d. analogue signal-processing applications. It permits each tap output to be individually accessed, thereby allowing programmable transversal filters and correlators to be implemented. Although the efficacy of the floating-gate tapping structure has been established, there is little information available describing the detailed operation of the tap circuit or presenting a suitable design procedure. It is the objective of this paper that information is presented relating all aspects of the sensing structure design, and to provide a basis for further definitive work. A list of symbols is provided in Appendix 14.1.

A diagram of a floating-gate tap circuit is given in Fig. 1. The c.c.d. tap electrode (or sense gate) is isolated during charge sensing, but, periodically, is reset to an intermediate voltage when there is no signal charge present underneath it. The 'floating' electrode is connected to the input gate of a sensing m.o.s.t., and any charge subsequently injected under this electrode stimulates a charge redistribution within the tap structure. This causes a related change in the electrode potential, which buffered by the m.o.s.t. tap circuit, to provide an output signal at a low impedance. The sensed-charge packet is then free to transfer along the delay line and the operating cycle is repeated. This modified clocking arrangement is referred to as 'pseudo-m-phase' operation for an $(m + 1)$ - phase c.c.d., e.g. pseudo $2 - \phi$, or $2\frac{1}{2} - \phi$ for a $3 - \phi$ device.

Although the floating-gate tap appears a simple circuit to design, careful consideration must be given to the resultant performance characteristics. Intrinsic c.c.d. parameters such as charge-transfer efficiency and charge-handling capacity are indeed impaired by the asymmetric-electrode structure. Additionally, the sense amplifier tends to dominate the circuit topology and restricts the system bandwidth.

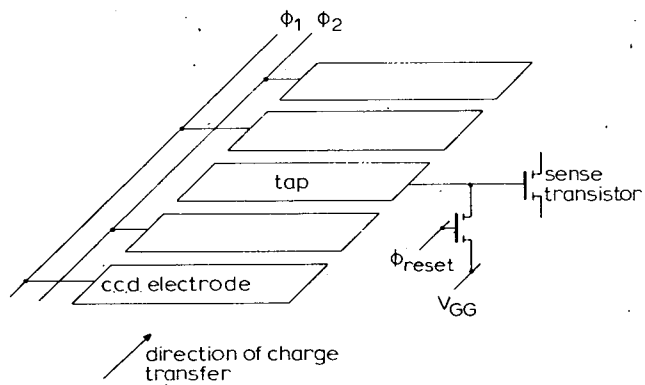


Fig. 1 *Floating-gate reset tap schematic for a 3-phase c.c.d. with pseudo-2-phase clocking*

In the following Sections, therefore, all aspects of the consequence of floating-gate tapping are considered, some relating to the c.c.d. operation, and others concerning the tap circuit design.

2 Fundamentals

A capacitive model of the tap structure for a surface-channel c.c.d. is shown in Fig. 2, where the total depletion and oxide capacitances under the tap electrode are given by C_D

Paper T67S, first received 8th April and in revised form 3rd May 1977

Dr. Deyner is with the Wolfson Microelectronics Liaison Unit, University of Edinburgh, Edinburgh, Scotland, and Dr. Mavor is with the Electrical Engineering Department, University of Edinburgh, Edinburgh, Scotland

and C_O, respectively, and C_L represents the total external load capacitance on the tap.*

A rigorous solution for the change in tap potential C_{sig} upon the introduction of a charge packet Q_{sig} is complicated by the variation of depletion capacitance with surface potential ϕ_s. However, such a solution has been obtained by MacLennan,[1] who considered the simultaneous solution of the following characteristic equations:

$$V_{sig} = Q_{sig}C_O/(C_OC_D + C_DC_L + C_OC_L) \quad (1)$$

$$C_D = C_0[V_0/2\phi_s]^{1/2} \quad (2)$$

$$\phi_s = V_0 + V_G - V_{FB} + Q_{sig}/C_0 - [2(V_G - V_{FB} + \frac{Q_{sig}}{C_0} V_0 + V_0^2]^{1/2} \quad (3)$$

where A is the tap-gate area, V_0 is a process-dependent constant given by $(qN\epsilon_0\epsilon_s)/C_{ox}^2$, and the tap-gate potential V_G is taken as the sum of a reset component V_{GG} and a signal component V_{sig}. The solution gives a 6th-order polynomial in V_G that may be solved numerically to give the transfer function shown in Fig. 3. Also shown is the limiting sensitivity relationship

$$V_{sig} = \frac{Q_{sig}}{C_L} \quad (4)$$

which is a good approximation to eqn. 1 for $C_D \ll C_O$ or $C_D \ll C_L$. The inequalities are achieved practically when the substrate doping N is small, or when a suitable substrate bias is applied.

3 System linearity

Although the intrinsic tap-transfer characteristic is non-linear, an overall linear-voltage transfer function may be obtained by employing a suitable input technique. Consider initially a tap structure with a constant-load capacitance C_L. It can be seen from the capacitive model of Fig. 2 that a linear relationship exists between any stimulated change in surface potential and the corresponding signal appearing at the gate, such that

$$V_{sig} = \frac{\Delta\phi_s C_O}{C_O + C_L} \quad (5)$$

where $\Delta\phi_s$ is the change in surface potential with respect to the 'empty well' level. By using an input method that sets the surface potential directly (e.g. diode cutoff[2]) and an f.g.r. tap structure in place of the metering well, as shown on Fig. 4, it is thus possible to obtain a linear input/output relationship at the metering tap. The amount of charge trapped will not be a linear function of the signal, but, when transferred along the register, it will give rise to identical linear-output signals in taps that are an exact replica of the metering tap. (For this reason, and because clock breakthrough may differ at the metering tap from that present at other taps, it is preferable that a separately adjustable reset level is provided for the metering tap, so that all reset potentials are identical after breakthrough.)

In general, the load capacitance C_L may not be sufficiently constant to permit the linearity required. Furthermore, the sense transistor of Fig. 1 may form part of a non-linear buffer state so that the usable output signal is a distorted version of the tap-gate signal. In such cases, it is possible to linearise the overall transfer function by employing an additional amplifier at the input. Essential to these techniques is the inclusion of a metering-tap circuit within the c.c.d. input structures; subsequent taps along the c.c.d. register should be electrical replicas of this tap, and will thus duplicate the metered output for a given signal-charge

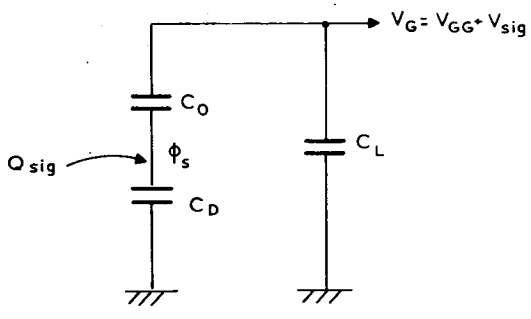


Fig. 2 *Capacitive model of floating-gate tap*

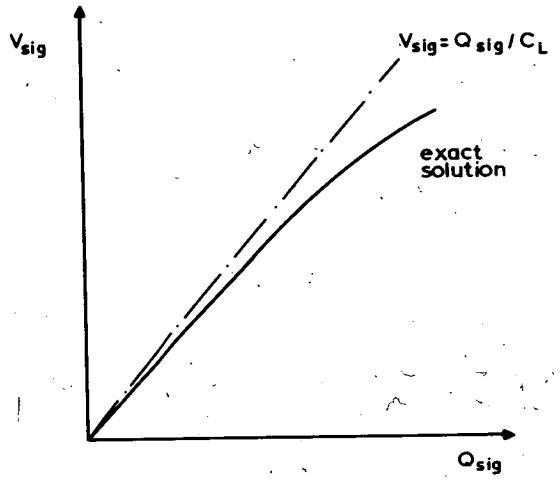


Fig. 3 *Variation of tap potential with injected charge*

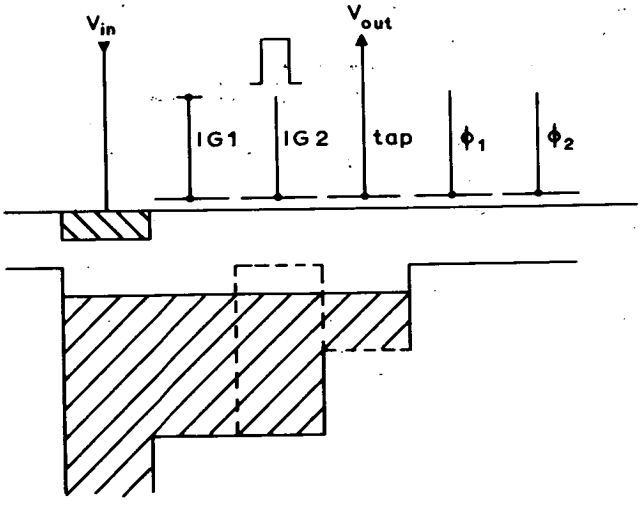


Fig. 4 *Linear input structure including input-potential metering tap*

*Throughout the paper, upper-case subscripts are used to denote absolute capacitive values, whereas lower-case subscripts indicate capacitance per unit area.

packet. By monitoring the metered output and regulating the injected charge before launching, the system transfer function may be linearised, regardless of the linearity of the intrinsic c.c.d. transfer function. In addition, the system transfer function may be constricted to unity gain, which is of great use when cascading devices.

One such technique, proposed by McLennan,[3] involves the inclusion of the nonlinear c.c.d. input system within the negative-feedback loop of a high-gain differential amplifier, as shown in Fig. 5*a*. In this way the metered output is used to modulate the applied c.c.d. input so that the sensed-output signal is made equal to the system input. Thus the total harmonic distortion due to both the c.c.d. input and the differential amplifier is attenuated by the open-loop gain of the system. Clearly, the scheme is only beneficial where the net steady-state closed-loop harmonic distortion is less that that obtained with the intrinsic c.c.d. input structure; thus imposing certain restrictions where onchip amplifiers are to be designed. The maximum attainable sampling frequnecy depends upon the slew rate and settling time of the closed-loop system, which should be critically damped.

An alternative linearisation scheme, shown in Fig. 5*b*, employs a controlled input ramp that is interrupted by a fast-switching comparator when the metered output is judged to be equal to the defined input voltage. Such a technique is proposed by Hense and Collings,[4] where the input ramp is achieved automatically by dynamic-charge injection. The accuracy and speed of these schemes is limited primarily by the resolution of the comparator and the response of the sense amplifier.

Both techniques are applicable to surface- and buried-channel devices for sample rates up to 1 MHz, although the latter appears the more suitable for 2-phase structures. Other configurations are certainly possible, and further investigation of these techniques is being pursued, with due regard to monolithic integration of the additional amplifier.[5]

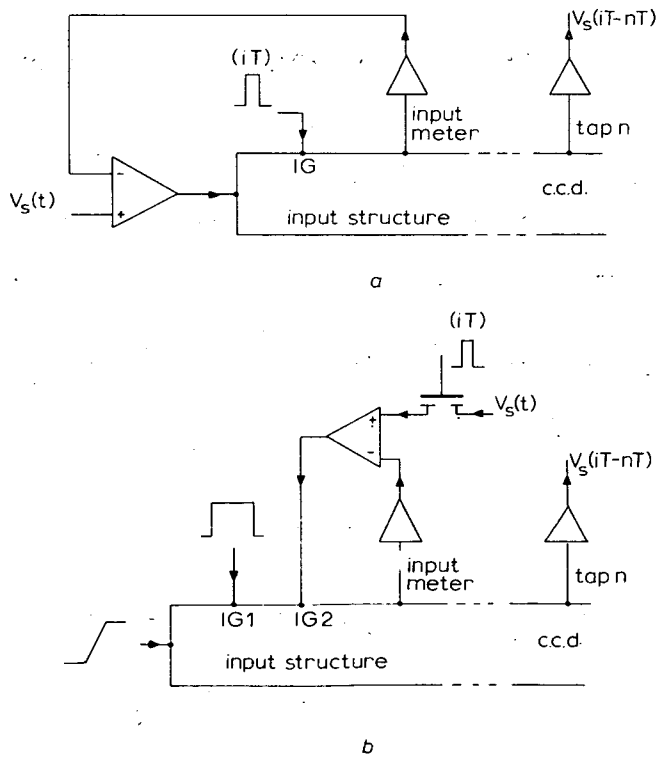


Fig. 5 *Two feedback linearisation schemes*

4 Charge-transfer efficiency

The presence of unlocked tap electrodes within the c.c.d. structure necessary to perform f.g.r. sensing tends to impair the overall transfer efficiency of the device at both low- and high-clock frequencies. This Section considers the effect of pseudo-2-phase operation on transfer efficiency and outlines procedures for optimising the design for either low- or high-frequency operation.

The charge-transfer efficiency of a surface-channel c.c.d. is limited at low-clock frequencies by interface state trapping at the semiconductor surface.[6] These states trap free minority carriers immediately, but release them with time constants dependent upon the state energies. At low clock frequencies the effect is such that carriers trapped from one charge packet are released into the following potential well. The amount of charge trapped is proportional to the surface area occupied by the charge packet, although the effect upon the signal information can be reduced by employing a small, but finite, background-charge level, or 'fat zero', which always fills the surface states and leaves the signal charge unaffected. However, the sides of the potential well formed under an electrode are not perfectly steep, and charge packets of differing amplitude occupy slightly different surface areas, as shown in Fig. 6. In this way, larger charge packets lose proportionally more charge. Assuming that the sides of the potential well slope linearly, the net effect is a constant transfer inefficiency, ϵ_e. Clearly, ϵ_e may be reduced by increasing the proportion of charge shielded by the fat zero, and by making the walls of the potential wall as steep as possible.

For the tapped register designs considered here the width/length ratio of individual electrodes is high, so that additional charge trapping occurs predominantly along the electrode edges adjacent to the interelectrode gaps, shown in cross-section in Fig. 6. We may conclude that the limiting low-frequency transfer inefficiency is inversely proportional to the electrode length L (which determines the surface area shielded by the fat zero) and to the electric fields present at either edge (which determine the slope of the sides of the potential well). Thus

$$\epsilon_e \alpha \frac{1}{L}\left(\frac{1}{V_{on}-V_{off_1}}+\frac{1}{V_{on}-V_{off_2}}\right) \qquad (6)$$

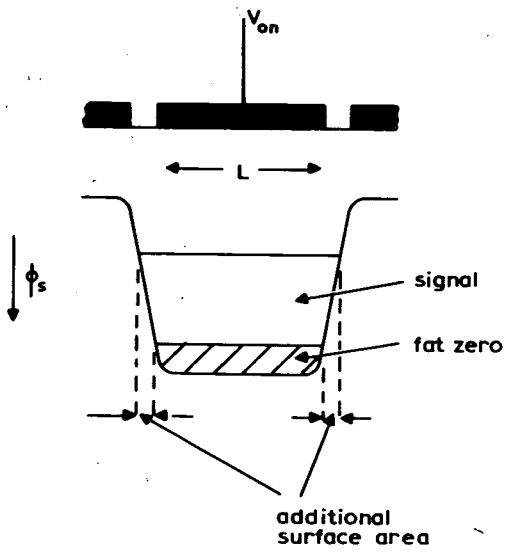


Fig. 6 *Edge effect of potential well wtih sloping walls*

where V_{on} is the 'on' potential of the electrode under consideration and $V_{off_{1,2}}$ are the 'off' potentials of the two adjacent electrodes.

Unfortunately, absolute evaluation of ϵ_e depends upon an estimate of the interelectrode fields by computer simulation of the structure. However, if the charge-transfer efficiency at low clock frequencies is of prime importance, then we may use eqn. 6 to choose optimum values of tap-reset potential and electrode lengths.

For small inefficiencies it may be assumed that the total charge-transfer inefficiency per cell is given by

$$\epsilon_e' = \sum_i^P \epsilon_i \tag{7}$$

where ϵ_i are the contributions from each of the P transfers in a given c.c.d. cell.

To simplify the analysis, we now define three types of electrode configuration that occur in floating- or biased-gate c.c.d. structures:

type α: electrode biased between two clocked electrodes
type β: clocked electrode between one biased electrode and one clocked electrode
type γ: clocked electrode between two clocked electrodes.

Now consider a general c.c.d. structure consisting of $N_{\alpha,\beta,\gamma}$ electrodes of lengths $L_{\alpha,\beta,\gamma}$, respectively. Eqns. 6 and 7 may be used to generate an expression for the limiting charge-transfer inefficiency ϵ_e' of the complete structure

$$\epsilon_e' = h\frac{N_\alpha}{L_\alpha}\frac{2}{V_{GG}} + \frac{N_\beta}{L_\beta}\frac{1}{\hat{V}_\phi - V_{GG}} + \frac{1}{\hat{V}_\phi} + \frac{N_\gamma}{L_\gamma}\frac{2}{\hat{V}_\phi} \tag{8}$$

where h is a constant and $\hat{V}_\phi$ is the peak clock potential and V_{GG} is the tap-reset potential, measured with respect to the minimum clock potential.

Eqn. 8 assumes that the edges of the potential well under the biased electrode slope linearly. This is strictly true for constant bias applications only. When the signal charge is used to modulate the biased potential, as in f.g.r. tapping, the surface area covered by the charge packet is no longer a linear function of its magnitude. The analysis is valid, however, if we assume that the average slopes of the sides of the potential well are equal to those corresponding to the fixed-bias case.

It is shown in Appendix 14.2 that, for a given total length restriction, ϵ_e' is minimised for

$$L_\alpha : L_\beta : L_\gamma = 1 : 1 : 1/\sqrt{2} \tag{9}$$

and

$$V_{GG} = \frac{\hat{V}_\phi}{2} \tag{10}$$

These expressions may be used to optimise the low-frequency performance of the c.c.d. register, regardless of the particular process parameters which determine the absolute value of ϵ_e'.

Pseudo 2-phase operation also adversely affects the high-frequency performance of the device where free-charge-transport mechanisms limit transfer efficiency. In particular, the time available for charge transfer with biased-gate operation can be considerably less than that available with conventional fully-clocked operation. Considering the three relevant transfers in detail with reference to Fig. 7, we note that:

(*a*) The phase-to-phase transfer takes place in the conventional manner, initially, but when the clock- and tap-surface potentials become equal, any remaining charge tends to partition equally between the tap well and the well being charged, as there is nothing to prevent the backward flow of charge underneath the tap electrode. The time available for transfer t_{a_1} is thus the time taken for the clock voltage to reach the tap-reset potential. This transfer may be taken to occur with an effective clock driving potential of $(\hat{V}_\phi - V_{GG})$ volts.

(*b*) The phase-to-phase transfer does not commence until the surface potential under the clock electrode has reduced to that under the tap electrode. Thereafter, the transfer may be assumed to take place conventionally, in time t_{a_2}, with an effective clock driving potential of V_{GG} volts.

(*c*) The tap-to-phase transfer commences immediately and can be assumed to be conventional during the period t_{a_3} with an effective clock driving potential of $(\hat{V}_\phi - V_{GG})$ volts.

For given driving conditions, therefore, the times available for transfer may be determined by studying the clock waveforms used. It can be seen from Fig. 6 that for typical pseudo-2-phase clock waveforms the time available for the phase-to-phase transfer t_{a_1} is notable restricted.

The complexity of free-charge transfer analyses prohibits an analytical optimisation of the f.g.r. structure with regard to high-frequency charge-transfer inefficiency. The following procedure does, however, allow the designer to specify ideal timing diagrams for a given c.c.d. structure, to achieve, for example, a target charge transfer inefficiency of 10^{-4} per transfer. The expressions are derived from two papers presented by Carnes *et al.*,[7, 8] who examined inefficiency as a function of transfer time for given gate lengths, clock voltages and process parameters.

A characteristic transit time τ_{tr} may be defined for a given electrode configuration, assuming infinite depletion width, as

$$\tau_{tr} = \frac{2L^3}{13 t_{ox} V_\phi^* \mu^*} \tag{11}$$

where V_ϕ^* is the effective clock driving potential (previously considered for the three cases of interest). Now an inefficiency of 10^{-4} per transfer requires an available transfer time[8] of $4\tau_{tr}$; thus we obtain an expression for the time that should be made available for transfer

$$t_a \simeq \left(\frac{L}{L_{cf}}\right)^3 \frac{2}{3\mu^* t_{ox} V_\phi^*} \tag{12}$$

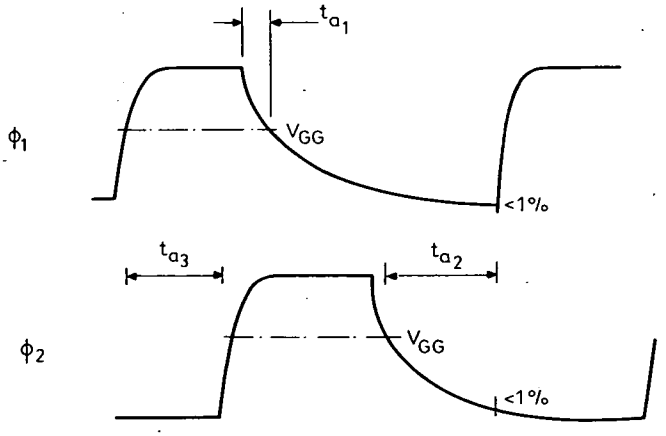


Fig. 7 *Typical floating-gate clock waveforms showing critical transfer times*

The length correction factor L_{cf} is included to account for the effect of finite depletion width x_d and is given by

$$L_{cf} = \left[\frac{1}{1 + \frac{1}{5}\frac{L}{x_d}}\right]^{4/3} \tag{13}$$

which is generally evaluated for half-full potential wells. These equations may be solved to obtain t_a for each electrode, thus completing the clock-driving specification in conjunction with Fig. 7. Clearly, one possible design procedure is to optimise the low-frequency charge-transfer efficiency and then use eqns. 12 and 13 to specify ideal clocking waveforms.

If the resultant clock waveforms are unsuitable, then it is possible to design electrode lengths to conform to the available transfer times. Eqn. 12 can be rearranged to define L in terms of t_a and used iteratively with a plot of L_{cf} against L from eqn. 13. This design procedure is not generally suitable, however, where the sense amplifier topology imposes a cell-pitch restriction, as considered in Section 8.

5 Signal handling

A basic parameter of tapped c.c.d. delay line design is the choice of output-signal magnitude. This is determined by the combination of tap sensitivity and peak-charge handling capability so that

$$\hat{V}_{sig} \simeq \frac{\hat{Q}_{sig}}{C_0} \tag{14}$$

where $\hat{Q}_{sig}$ is the maximum charge handling capability of the structure. Now the value of $\hat{Q}_{sig}$ is limited by the fastest-filling potential well, which may be associated with the tap electrode or either of the adjacent clock electrodes, depending on the tap-reset potential and relative electrode lengths. It is thus necessary to study the filling of each well. For this purpose a surface-potential diagram of the structure is given in Fig. 8, where all potentials are given with reference to the lowest clock potential and variation of the reset level due to clock breakthrough is ignored.

Now the surface potential under a clock electrode, $\phi_{s_{clk}}$, as a function of total injected charge Q_{sig} is given by

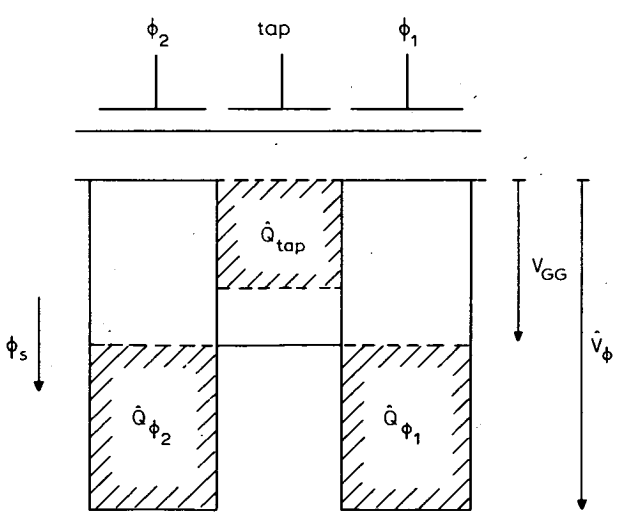


Fig. 8 *Surface potentials and charge-handling capabilities of floating-gate c.c.d. stage*

$$\phi_{s_{clk}} \simeq V_0 + \hat{V}_\phi - V_{FB} + \frac{Q_{sig}}{C_0} \tag{15}$$

from which it can be seen that the potential well fills linearly,[9] giving a charge capacity

$$\hat{Q}_{sig} \simeq W_S C_{O_{clk}}$$

where W_S is the maximum allowable change in surface potential. The value of W_S is given by the depth of the potential well $(V_{GG} - \hat{V}_\phi)$ so that

$$\hat{Q}_{clk} = (V_{GG} - \hat{V}_\phi)C_{0clk} \tag{16}$$

Consideration of the surface potential under a tap electrode,

$$\phi_S \simeq V_0 + V_{GG} + V_{sig} - V_{FB} + \frac{Q_{sig}}{C_{0tap}} \tag{17}$$

shows that this potential well fills faster than that under a normally clocked electrode, due to the interaction between the induced signal and the surface potential given by eqn. 5. Combining these equations and using the potential well depth $W_S = -V_{GG}$ from Fig. 5 we obtain

$$\hat{Q}_{tap} = -V_{GG}\frac{C_L C_{0tap}}{C_L + C_{0tap}} \tag{18}$$

By combining eqns. 16 and 18 with the sensitivity relationship, eqn. 14, the peak attainable signal voltage may be evaluated for the two cases of interest as follows:
(*a*) charge limited by the tap well,

$$\hat{V}_{sig} = \frac{-V_{GG}}{1 + \frac{C_L}{C_{0tap}}} \tag{19}$$

and
(*b*) charge handling limited by the clocked well,

$$\hat{V}_{sig} = -(\hat{V}_\phi - V_{GG})\frac{C_{0clk}}{C_L} \tag{20}$$

where the negative signs occur because the signal appears as a reduction in the tap potential.

Clearly, for a given reset level, clock amplitude and known electrode lengths, the ouput-signal magnitude may be determined by a suitable choice of capacitance ratio. The effects of background charge and reset-level breakthrough are to reduce the peak attainable output signal by some 10%, depending on their relative magnitudes.

6 Noise performance

To achieve a given output dynamic range it is necessary to consider all noise sources associated with charge input and transfer, and signal tapping and sensing. For convenience, the effect of each noise contribution is commonly expressed as an equivalent r.m.s. fluctuation in the number of carriers in each charge packet, and a comprehensive analysis in these terms has been given by Caranes and Kosonocky.[10] A summary of the individual noise contributions is given in Appendix 14.3.

For typical signal processing applications, the noise source of most concern is that of fluctuations in the occupancy of fast-interface states at the semiconductor sur-

face. The resultant r.m.s. carrier fluctuation N_{FIS} is given by[10]

$$N_{FIS} = [1{\cdot}4\,kTN_{SS}N_gA_g]^{1/2} \quad (21)$$

where N_{SS} is the surface-state density, N_g is the number of transfers, and A_g is the average gate area. It can be seen that for any given process and register length this noise term is proportional to the square root of the gate area. Now for a given output signal level the peak charge capacity increases linearly with gate area provided the capacitive ratios derived in Section 5 are maintained. Thus, it is possible to increase the width of the c.c.d. channel (and the value of C_L correspondingly) to improve the signal/noise ratio to the desired level.

7 Dark current

Any f.g.r. tapping arrangement normally requires some on-chip sensing and buffering circuitry, which necessarily dissipates power. As the intrinsic c.c.d. structure is essentially nondissipating, the introduction of such a tapping system may adversely effect certain operating parameters through the associated increase in device temperature. Indeed, the dark current effect is known to be strongly temperature dependent.

Dark current is the name given to the flux of thermally generated minority carriers into the signal-charge packet. As these carriers are indistinguishable from signal charge, an additive error will result, and, because of the finite signal capacity, the dynamic range of the device may be impaired. Distortion of the signal is not as significant because the injection mechanism is essentially linear. Dark current arises mainly from electron-hole pair generation at the semiconductor-oxide interface and within the depletion region surround it.[11] The generation rate due to both sources is proportional to the intrinsic minority-carrier concentration

$$n_i \propto \exp(-E_a/kT) \quad (22)$$

where the activation energy E_a, for silicon, is taken as half the bandgap in electron volts. At normal operating temperatures the dark current may be expected to approximately double for every 10°C rise in temperature. The device temperature is related to the onchip power dissipation P and the package thermal resistance R by

$$T_c = T_a + RP \quad (23)$$

Thus, knowing the dark-current level at some ambient temperature, the device designer can predict the dark-current density J_D at a given power-dissipation level. The resultant additive charge Q_D after an integration time t_i, is given by

$$Q_D = J_D A_c t_i \quad (24)$$

where A_c is the average c.c.d. cell area. This charge level may be related to the peak-signal capacity of the c.c.d. to determine the resultant signal error or the reduction in available dynamic range.

8 Reset transistor

Ideally, the tapping mode is only capacitively coupled to ground so that the output signal should appear as a change in tap potential, with respect to a given initial reset level. In practice, some leakage current to ground is unavoidable, and it becomes necessary to continuously bias the tap gate, either by using a resistor 'pull-up' or by periodic resetting using a m.o.s.t. switch.

The resistor pull-up technique has the disadvantage of imposing a highpass filter characteristic on the tap so that signal information at low frequencies, and in particular at d.c., is lost. Additionally, it is not practical to realise monolithically the high values of resistance desirable for such a biasing technique.

The reset m.o.s.t. technique, included in the schematic of Fig. 1, preserves all of the signal information by periodically resetting the tap node to the desired potential when there is no signal charge present under the electrode. This may be achieved most conveniently with $3-\phi$ devices by resetting the node once every clock period, using the ϕ_1 clock waveform. A particular advantage of this technique is that the reset m.o.s.t. may be fabricated easily as part of the tap structure. As the tap nodal capacitance is small (typically 1 pF) a minimum-geometry reset transistor is found to be adequate for most applications. An improvement in the tap-output signal may be obtained by including an additional shield gate between the reset gate and the tap node, thus reducing capacitive breakthrough of the reset waveform once the transistor has turned off.

9 Sense amplifier

As the tap-output signal is intrinsically high impedance, a sense/buffer amplifier is commonly employed to provide a more useful low-impedance output. Such amplifiers may be implemented conveniently using m.o.s. transistors that are process-compatible with the c.c.d., and may be fabricated monolithically with it. They have the necessary property of essentially infinite input resistance and are capable of output admittances up to 1000 μS. Where system linearity is achieved with a feedback-input technique, the amplifiers should be well matched and should ideally preserve or enhance the useable signal magnitude without adversely affecting dynamic range. Careful study of various sense amplifier configurations suggests that these qualities may be best met by the m.o.s.t. source follower with constant-current load, shown in Fig. 9, and analysed in part by Frohman-Bentchkowsky *et al.*[12] It has the important advantage that the small signal gain is close to unity, highly linear and essentially independent of transistor matching. Threshold voltage and gain variations do, however, result in random d.c. offsets at the output, which, with good design and processing, may be limited to the order of ±50 mV.

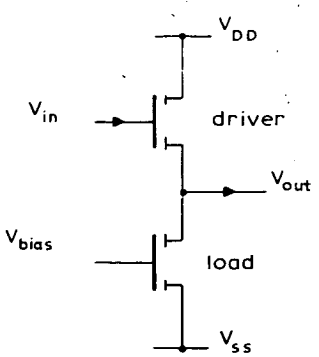


Fig. 9 *M.O.S.T. source follower with saturated constant-current load*

$V_{bias} < V_{out} + V_T$

The output conductance of the source follower Y_0 is given approximately by the transconductance of the saturated active tansistor or

$$Y_0 \simeq g'_m \simeq \sqrt{\left(2\beta_0 \frac{Z}{L} I_D\right)} \qquad (25)$$

Thus the dimensions of the sense transistor will normally be determined by the specified device parameters of output conductance and supply current (or power dissipation). The load device and bias voltage should be chosen to satisfy I_D in the most convenient way. For depletion-mode transistors, V_{bias} may be at ground potential and a suitably large load transistor used. Where many sense amplifiers are required, however, the layout area may be reduced by using minimum-dimension load transistors and generating the required bias voltage on chip. With good design the latter technique has the better tolerance to threshold-voltage variations.

At this point in the design, the geometry of the sense amplifier and reset-transistor structure may be specified; it should be noted here that in continuously tapped delay-line designs the geometry of the sense amplifier commonly determines the cell pitch. Thus the transfer efficiency considerations of Section 4 should be combined with this pitch restriction to determine individual electrode lengths. In many cases the two considerations are incompatible, and it may be necessary to employ two or more c.c.d. bits per sense amplifier cell in order to obtain the required high-frequency performance. The redundancy introduced by this technique may, in fact, be put to advantage[13] by multiplexing the signal input in such a way as to reduce the apparent transfer inefficiency.

Having designed the sense amplifier, its input capacitance C_{in} may be determined from a small signal model as

$$C_{in} = C_{gb} + C_{gd_e} + (C_{gs} + C_{gs_e})(1 - G_V) \qquad (26)$$

where G_V is the small signal gain (typically 0·9) and the subscript e indicates an external (overlap) capacitance. Each of the components of eqn. 26 is a function of the quiescent operating conditions (see for example Cobbold,[14] chap. 8) and their evaluation is somewhat laborious. A good design approximation is, however, given by

$$C_{in} \simeq C'_0 \cdot 1 - \frac{G_V}{2} \qquad (27)$$

where C'_0 is the total oxide capacitance associated with the active transistor.

The input capacitance of the sense amplifier is, of course, an integral part of the total tap load capacitance C_L, which may now be estimated by inspection of the layout and use of eqn. 27. From Section 5, the ratio of C_L to a particular electrode oxide capacitance defines the achievable output-signal magnitude. Having already determined the tap electrode length, this ratio may be used to define a channel width Z.

The intrinsic dynamic range of the structure may now be estimated from a knowledge of the peak-charge capacity (Section 5) and expected noise level (Section 6). If necessary, the c.c.d. channel width Z and the tap load capacitance C_L may be increased to obtain the noise performance required.

10 Design example

To illustrate some of the properties and capabilities of the f.g.r. structures examined in this paper, consider the design of a tapped c.c.d. register subject to the target specification given in Table 1. A complete circuit schematic of a single tap based upon the previous considerations is shown in Fig. 10, which also indicates the major design parameters. Fig. 11 shows a possible layout geometry for one cell of the tapped delay line for a double aluminium s.e.t. process.[15]

An immediately noticeable feature of this design is the inclusion of two c.c.d. stages per sense amplifier, a necessary measure to shorten gate lengths and thus realise the specified high-frequency performance. Note that the maximum clock frequency required to maintain the specified delay between taps is now 4 MHz (refer to Table 1).

The sense amplifier layout in this case imposes a cell pitch of 56 μm that has been apportioned between the individual electrodes, as far as layout rules allow, in such a way as to minimise charge-transfer inefficiency at low frequencies, following the analysis of Section 4.

Free-charge-transfer considerations yield critical transfer times of $t_{a_1} \geqslant 22$ ns, $t_{a_2} \geqslant 40$ ns and $t_{a_3} \geqslant 68$ ns for the tapped c.c.d. stage, where these times are indicated in Fig. 6. Such transfer times are adequate for the fully clocked c.c.d. stage because of the increased effective driving voltages and shorter electrode lengths. It is easily verified that clock waveforms conforming to these requirements may be generated at frequencies up to 4 MHz.

A reset level (after clock breakthrough) of half the peak clock voltage has been chosen. This allows a theoretical peak output signal of 4·5 V for the tap and load capacitances indicated, and allowing for a small-signal sense amplifier gain of 0·9. The total sensitivity is 1·34 V/pC. In practice, the maximum output signal may be reduced to the specified 4·0 V after accounting for a background 'fat zero' charge level and capacitive clock breakthrough on to the isolated tap electrode.

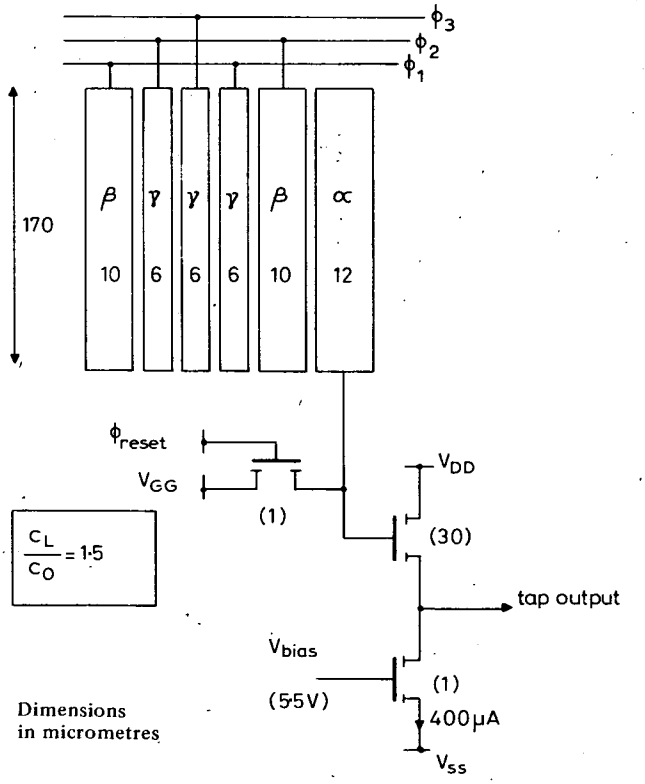


Fig. 10 *Schematic of single tap circuit showing electrode types*

The expected peak signal charge is 3·0 pC, or $2{\cdot}82 \times 10^7$ carriers. The theoretical noise level (Section 6) at the last tap is $4{\cdot}27 \times 10^3$ r.m.s. electrons, indicating that the specified dynamic range should be achieved so that no further increase in channel width or load capacitance is necessary.

When mounted in a ceramic package of thermal resistance 75°C/W, the chip temperature under operating conditions is expected to be 15°C above ambient. Thus, from Section 7, we may predict an increase in dark current of approximately three times the ambient value (in this case, 40 nA/cm^2). For the average bit area under consideration this results in an increase in the signal-charge packet of 2 pC/s. For a delay of 1 ms between taps, this dark-current level reduces the available dynamic range by up to 5% at the last tap. However, a dynamic range of 60 dB is still theoretically obtainable.

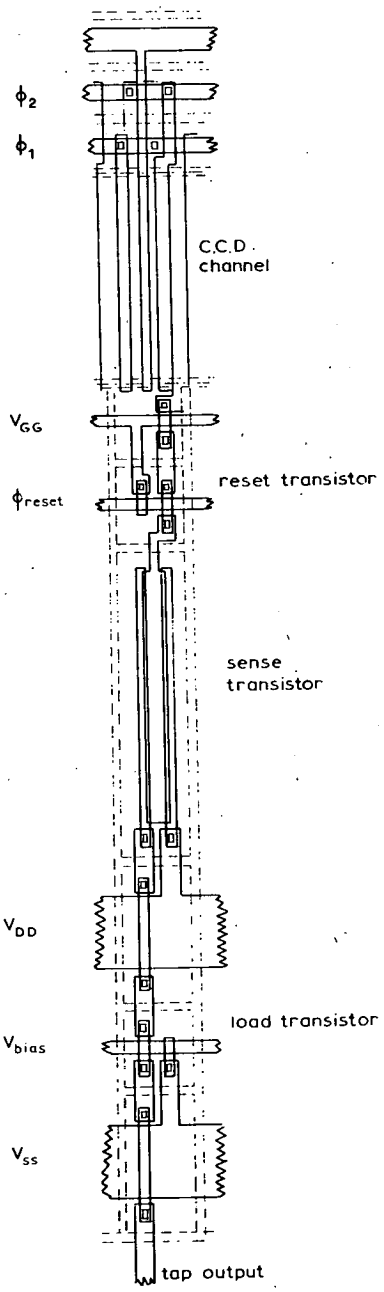


Fig. 11 *Typical layout of f.g.r. circuit (of Fig. 10)*

Table 1: Target specifiifation of a typical 32-tap c.c.d. delay line

Parameter	Value
Number of taps	32, plus input metering tap
Time delay	1 ms → 500 ns per tap
Sampling frequency	1 kHz → 2 MHz
Power dissipation	200 mW
Power supply	$V_{DD} = 15$ V, $V_{SS} = 0$ V
Tap output impedance	<2 kΩ
Clock magnitude	25 V peak-to-peak
Available output signal	4 V peak-to-peak
Dynamic range (at tap 32)	>60 dB
Total harmonic distortion	< −50 dB
With feedback linearisation	< −60 dB

11 Conclusions

Tapped c.c.d. delay lines using the floating-gate reset (f.g.r.) technique have been demonstrated in several analogue-signal-processing applications to date. The overall peformance of the tapped delay line depends on many factors, including the c.c.d. type and its gate dimensions, the limiting effect of the f.g.r. tapping cirucitry, the bias and clock voltages, the input technique, the starting material and the processing conditions etc.

This paper has presented a basic quantiative approach for the device engineer to undertake detailed and systematic design. The interaction between the various physical processes and circuit conditions that govern the device performance have also been reviewed. The description of the design example and the accompanying analysis should also help with the layout of the structure and indicate the geometrical limitations.

Although this work should aid the c.c.d. design engineer, it is by no means an exhaustive treatment, nor are all of the expressions optimised. Further analysis, device design and characteristisation will be necessary before a definitive design guide can be produced.

12 Acknowledgments

The authors acknowledge valuable work by D.J. MacLennan, discussions with G.F. Vanstone and assistance from the c.c.d. group at RSRE (Malvern) and the Wolfson Microelectronics Liaison Unit at Edinburgh. The research has been sponsored by the UK Science Research Council, DCVD MOD(PE) and AUWE, Portland.

13 References

1 MACLENNAN, D.J., MAVOR, J., VANSTONE, G.F., and WINDLE, D.J': 'Novel tapping technique for charge-coupled devices', *Electron. Lett.*, 1973, 9, pp. 610–611
2 SEQUIN, C.H., and MOHSEN, A.M.: 'Linearity of electrical charge injection into charge-coupled devices' (IEDM Technical Digest, Washington, D.C., 1974) pp. 229–232
3 MACLENNAN, D.J., and MAVOR, J.: 'Novel technique for the linerisation of charge-coupled devices', *Electron. Lett.*, 1975, **11**, pp. 222–223
4 HENSE, K.R., and COLLINS' T.W.: 'Linear charge-coupled device signal processing techniques', *IEEE J. Solid-state Circuits*, 1976, **SC-11**, pp. 197–202
5 WESTE, N., and MAVOR, J.: 'An MOS amplifier for c.c.d. applications'. Proceedings of c.c.d. technology and applications conference, Edinburgh, 1976, pp. 326–328
6 TOMPSETT, M.F.: 'The quantitative effects of interface states on the performance of charge-coupled devices', *IEEE Trans.* 1973, **ED-20**, pp. 45–55
7 CARNES, J.E., KOSONOCKY, W.F., and RAMBERG, E.G.: 'Drift-aiding fringing fields in charge-coupled devices', *IEEE J. Solid-State Circuits*, 1971, **SC-6**, pp. 322–326

8 CARNES, J.E., KOSONOCKY, W.F., and RAMBERG, E.G.: 'Free charge transfer in charge-coupled devices', *IEEE Trans.*, 1972, **ED-19**, pp. 798–808
9 BEYNON, J.D.E.: 'The basic principles of charge-coupled devices', *Microelectronics*, 1975, **7**, pp. 7–13
10 CARNES, J.E., and KOSONOCKY, W.F.: 'Noise sources in charge-coupled devices', *RCA Rev.*, 1972, **33**, pp. 327–343
11 TASCH, A.F., BRODERSEN, R.W., BUSS, D.D., and BATE, R.T.: 'Dark current and storage-time considerations in charge-coupled devices'. Proceedings of the c.c.d. applications conference, San Diego, 1973
12 FROHMAN-BENTCHKOWSKY, D., and VADASZ, L.: 'DC analysis of an MOS source follower', *IEEE J. Solid-State Circuits*, 1968, **SC-3**, pp. 306–307
13 MAVOR, J., DAVIE, M.C., and DENYER, P.B.: 'Techniques for increasing the effective charge-transfer efficiency of tapped c.c.d. registers', *Electron. Lett.*, 1977, **13**, pp. 31–33
14 COBBOLD, R.S.C.: 'Theory and applications of field-effect transistors', (Wiley, 1970)
15 PERKINS, K.D., and BROWNE, V.A.: 'Sub-micron gap metal gate technology for CCD's', *Microelectronics*, 1975, **7**, pp. 14–22

14 Appendixes

14.1 List of symbols

A_g = average c.c.d. gate area
A_c = average c.c.d. cell area
C_D = total depletion capacitance under c.c.d. electrode
C_L = total tap load capacitance
C_0 = total oxide capacitance under c.c.d. electrode
C_0^1 = total capacitance associated with sense transistor
C_{ox} = oxide capacitance per unit area
g_m = small-signal transconductance
I_D = source-follower drain current
J_D = dark-current density
k = Boltzmann's constant
L = c.c.d. gate length
n_i = intrinsic carrier concentration
N = substrate doping concentration
q = electronic charge
$\hat{Q}_{clk}$ = peak charge capacity under clocked electrode
Q_{sig} = signal charge
$\hat{Q}_{tap}$ = peak charge capacity under tap electrode
t_a = time available for charge transfer
t_{ox} = oxide thickness
T = absolute temperature
T_a = ambient temperature
T_c = chip temperature
V_{FB} = m.o.s. flat-band voltage
V_G = tap-gate potential
V_{GG} = tap-reset potential (after clock breakthrough)
V_0 = process dependent constant
V_{sig} = signal potential at tap gate
V_T = transistor threshold voltage
V_ϕ = clock voltage
V_ϕ^* = effective driving potential for charge transfer
$\hat{V}_\phi$ = peak clock voltage
x_d = depletion-region width
Z = c.c.d. channel width
β_0 = transistor gain factor
ϵ_e = charge transfer inefficiency due to the 'edge effect'
ϵ_e' = transfer inefficiency per c.c.d. cell
ϵ_0 = absolute permittivity of free space
ϵ_s = relative permittivity of silicon
μ^* = effective carrier mobility
τ_{tr} = characteristic charge transfer time constant
ϕ_s = surface potential
$\Delta\phi_s$ = change in surface potential

14.2 Optimisation of low-frequency transfer inefficiency

Let the electrode lengths be combined in the ratios

$$K_1 = \frac{L_\beta}{L_\alpha} \quad \text{and} \quad K_2 = \frac{L_\gamma}{L_\alpha} \tag{28}$$

and impose as a boundary condition a total structure length L':

$$L' = N_\alpha L_\alpha + N_\beta L_\beta + N_\gamma L_\gamma \tag{29}$$

Equation 8 may be rewritten as

$$\epsilon_e' = \frac{h}{L'}(N_\alpha + K_1 N_\beta + K_2 N_\gamma)\left[\frac{2N_\alpha}{V_{GG}} + \frac{N_\beta}{K_1(2\hat{V}_\phi - V_{GG})} + \frac{2N_\gamma}{\hat{V}_\phi}\right] \tag{30}$$

To minimise ϵ_e' we must choose K_1, K_2 and V_{GG} so that

$$\frac{\partial \epsilon_e}{\partial K_1} = \frac{\partial \epsilon_e'}{\partial K_2} = \frac{\partial \epsilon_e'}{\partial V_{GG}} = 0 \tag{31}$$

Evaluating the partial derivatives and solving simultaneously

$$L_\alpha^2 : L_\beta^2 : L_\gamma^2 = 2:2:1 \tag{32}$$

$$V_{GG} = \frac{V_\phi}{2} \tag{33}$$

14.3 Noise sources[10]

The dominant noise sources in charge-coupled devices are broadly of two types: those associated with charge transfer, which tend to be cumulative, and those associated with charge injection and sensing, which are purely additive.

One source associated with charge transfer arises from fluctuations in charge-transfer efficiency so that

$$\bar{N}_{n_\epsilon} = [2\epsilon N_g N_s]^{1/2} \tag{34}$$

where N_s is the total number of carriers per charge packet including background ('flat zero') charge

A more dominant cumulative noise source, that is independent of signal charge, is attributable to fluctuations in the total number of carriers trapped by fast interface states so that

$$\bar{N}_{n_{FIS}} = [1{\cdot}4\, kTN_{ss}N_g\bar{A}_g]^{1/2} \tag{35}$$

where N_{ss} is the fast-interface state density and $\bar{A}_g$ is the average gate area.

Noise due to charge setting at the input and due to sense node resetting is characterised by the capacitance of the node being charged, so that

$$\bar{N}_{n_{reset}} = 400(C_{pF})^{1/2} \tag{36}$$

where C_{pF} is the nodal capacitance in picofarads.

Finally, noise sources within the m.o.s.t. sense amplifier must be considered. Noise models tend to be complex, but for most applications this source is small compared to those already considered. An approximate expression, obtained by Carnes and Kosonocky, is

$$\bar{N}_{n_{most}} = 60C_{pF}\left[\frac{\Delta B}{5\ \text{MHz}} \times \frac{1000\,\mu S}{g_{mo}}\right]^{1/2} \tag{37}$$

where C_{pF} is the sense nodal capacitance in picofarads, ΔB is the filtered output bandwidth and g_{mo} is the transconductance of the sense transistor in microsiemens

Design and development of c.c.d. programmable transversal filters

J. Mavor and P.B. Denyer

Indexing terms: Charge-coupled devices, Correlators, Filters, Signal processing

Abstract: This paper examines the potential of integrated-circuit-based programmable transversal filters for baseband operation. The emphasis is on compact processors with an eventual aim of a single chip, monolithic design suitable for a variety of signal-processing applications. Before hardware developments are treated, trade-offs between serial and parallel configurations of convolver are considered, together with their respective component requirements. It appears that for complete integration of a transversal filter module, the parallel approach is preferred. A study is made of the circuit elements of a particular realisation of an integrated analogue convolver in c.c.d./m.o.s. technology. All aspects of the design are considered and the description is focused around a prototype 64-stage programmable transversal filter. Several such units could be cascaded giving this integrated approach an enormous potential in applications where space and power are crucial, and high computing precision is not essential.

1 Introduction

The development of compact, lightweight, low-power hardware to perform the real-time correlation/convolution of two discrete signals permits the realisation of sophisticated signal processing functions, such as matched filtering and spectrum analysis for a variety of applications. Recent advances in microelectronic components are central to these developments and include progress in analogue as well as digital devices. Now, in addition to conventional digital elements for performing discrete-time signal processing, charge-coupled devices[1] (c.c.d.s) are available. These provide primarily high-density analogue/digital storage in various configurations, and are compatible with m.o.s. circuitry for performing ancillary functions. The c.c.d/m.o.s. approach looks extremely promising for realising convolvers with performance attributes suitable for many areas including sonar, communications and instrumentation, particularly for mobile equipments.

The advent of the microprocessor and l.s.i. memory means that powerful and versatile filtering systems can be programmed and controlled. Such an arrangement is extremely important in the signal-processing context as any desired filtering function can be realised simply by programming its impulse response. Interfaced with a microprocessor[2] this system could provide 'intelligent' adaption of the filter response to optimise performance, or, for restricted applications, a single filter may be time multiplexed to implement many different signal processing functions.

A programmable transversal filter can be realised using various architectures based upon combinations of serial and parallel forms. Each configuration is likely to be suited to a particular hardware arrangement and dependent upon the overall system parameters. In this paper, the emphasis is on solid-state convolvers based on silicon integrated circuit technology. The trade-offs between, and limitations of, serial and parallel forms of processor are discussed with particular reference to integrated circuit implementation with a minimum package count.

Paper T110 E, First received 1st September and in revised form 8th November 1977

Dr. Mavor is with the Electrical Engineering Department and Mr. Denyer is with the Wolfson Microelectronics Liaison Unit, University of Edinburgh, Mayfield Road, Edinburgh EH9 3JL, Scotland

The paper is based on a paper presented at the AGARD/NATO meeting on the impact of charge-coupled devices and surface-acoustic waves on signal processing and imagery in advanced systems, Ottawa, Canada, 11–15th October 1977. AGARD Report 230

2 Principles

2.1 Convolution

A filter having an impulse-response sequence r of N elements will respond to an L element signal sequence s with an $(L+N)$ element ouput sequence given by the convolution sum[3]

$$(s*r)(m) = \sum_{n=0}^{N-1} r_n s_{m-n} \quad (1)$$

The calculation requires N multiplications for each convolved point (one multiplication for each reference element) and, thus, a total of $(L+N)N$ multiplications for the complete convolution sequence.

2.2 Serial versus parallel

Large computers can be used to program and carry out these computations very easily, but the demand for dedicated, compact, hardware realisations is an important and developing applications area.

All convolution systems must provide some memory register for storing the reference sequence and the last N signal samples, $s_{m-n} \ldots s_m$, a facility for multiplying together individual samples of the two sequences, and a method of shifting the sequences with respect to each other, by one element, after each calculation of the convolution sum.

One common approach for realising the convolution function is shown in principle in Fig. 1*a*, and employs a single multiplier to execute the N multiplications at each convolved point in *series*. Thus for real-time convolution, the multiplier must operate at N times the signal sample frequency; similarly, the data contained in the signal and reference registers must be recirculated at that rate. After each recirculation, the oldest signal sample is replaced by a new input sample and the next convolution point is obtained by a further recirculation, with N corresponding multiplications in series.

As an alternative to the serial system, real-time computation of the convolution sum may be achieved at a

0308–6984/78/110E–0001 $1·50/0

processing rate equal to the signal sampling rate by the effective parallel processing of the required multiplications. This may be realised in an N-stage, programmable transversal filter of the type shown in Fig. 1*b*. Here the hardware requirement is increased by N-1 multipliers, but for a given signal bandwidth the speed requirement is reduced by a factor of N. Note also that the equivalent number of signal sample-shifts, or transfers, is also reduced by a factor N.

Many other correlator architectures are possible, usually involving one or more serial/parallel configurations to optimise the parameters of the system. In many cases the speed of the multiplication hardware limits the operating bandwidth and the system has to be configured with multiplexed or parallel channels to achieve performance at the expense of duplicated multiplication.

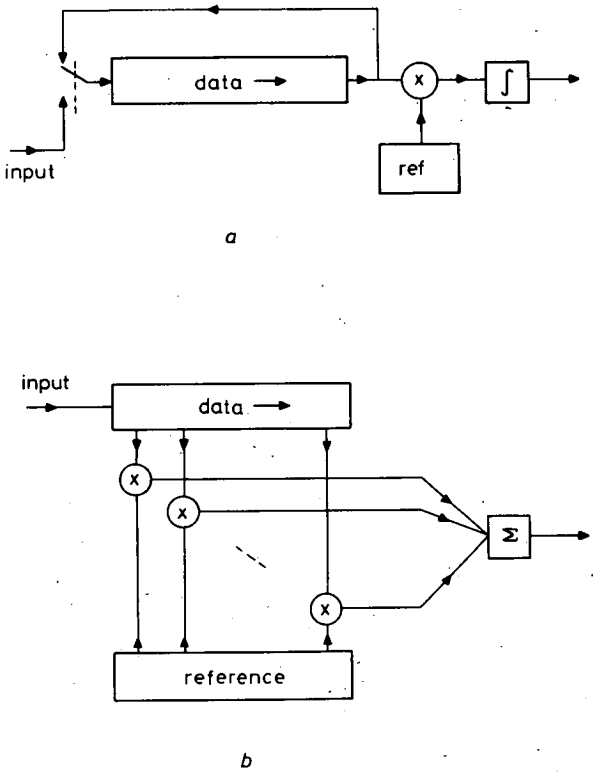


Fig. 1 *Serial and parallel forms of implementing the convolution sum*

2.3 Analogue versus digital processing

Convolvers based on conventional digital components are widely available but suffer from the disadvantage that when they are used to process analogue signals – as is very often the case in signal processing – a./d. conversion is required. Their speed is limited essentially by the rate of the digital multipliers used, and these are likely to be largely responsible for the high power taken by such a processor. Furthermore, digital multiplication is costly in terms of hardware, requiring currently at least one package per multiplier. For higher packing density, analogue multipliers and summers can be used but these, of course, suffer from the inaccuracies of analogue multiplication and thus impose an upper limit on useful resolution. However, the distinct advantage of a digital convolver – in common with all digital systems – is that the accuracy of computation, which is a direct function of the digital word size, is essentially infinitely expendable (but naturally limited by the components used). This precision is obtained at the expense of power and system size, but for some applications is of major importance.

Analogue processors of all kinds suffer from the disadvantages of drift, sensitivity and difficult circuit design. Convolver realisations based on analogue techniques are also subject to the restrictions of the usually modest dynamic range of analogue-analogue signal multipliers. However, where these problems can be contained they have the attributes of a *more direct* realisation of the convolution function, in comparison with digital approaches; thus advantages may arise which outweigh the inherent problems for a particular application.

In realising an analogue processor, analogue memory could be provided by m.o.s. capacitor storage; however, by using charge-coupled-device memory then the shift facility is provided inherently. Thus, c.c.d.-based analogue convolvers offer a potential size advantage over all-digital techniques. Their application is substantially limited to parallel, or transversal-filter, realisations of convolver because of the difficulty of recirculating sampled analogue data in an analogue-mode c.c.d. Charge-transfer efficiency problems degrade the data after many recirculations and made serial convolver forms essentially impossible to implement, even with compensation for charge dispersion.

The provision of an analogue reference memory is an additional problem to be faced; conventional c.c.d./m.o.s. processes do not allow long-term storage of analogue information, and some means of refreshing must be provided. This may not be such a severe disadvantage at the systems level, however, as the reference information may often be held in digital form in some memory associated with a central control unit (possibly a microprocessor). Refreshing the reference can then take on a 'direct memory access' (d.m.a.) priority without interfering with other control functions. Thus a dedicated non-volatile memory block, always associated with a given analogue convolver, is not necessarily implicated.

A further convolver formation is to have an analogue signal input and a digital reference. The multiplier requirement is relaxed where a single bit reference is suitable and, consequently, the overall hardware complexity of the processor thereby becomes reduced; analogue multipliers are, however, implied. Monolithic correlators of this type have been reported[4] in c.c.d./m.o.s. technology. In order to simulate a multibit reference, however, several units must be run in parallel or one longer unit time-shared. Alternatively, a multibit reference may be used with multiplying d./a. convertors at each filter point. The complexity of this approach limits the filter packing density to 8 points per chip, currently; thus large time-bandwidth products are not readily fabricated.

2.4 Comparison of realisations

It is clear from the preceding discussion that if high precision is required then a digital processor is essential. For low bandwidth applications ($<$ 20 kHz for a time-bandwidth of 1000) then a serial convolver offers many advantages. However, where dynamic ranges of 40–60 dB are suitable, or a high speed processor is required, then a parallel approach based upon c.c.d./m.o.s. devices may be a possibility. Therefore, as digital component development progresses, the boundary between the various approaches will

Table 1: Performance comparison of different convolver realisations

Convolver type					
Configuration	Mode	Speed	Packing density	Accuracy	Features
Serial	digital	medium-slow (trade-off with packing density)	medium-high (trade-off with accuracy)	to any requirement (with digital multiplication)	permanent, signal and reference memory
	analogue and digital			limited by analogue multipliers	restricted by recirculation of analogue data
	analogue				
Parallel	digital	fast	low	to any requirement (with digital multiplication	permanent, signal and reference reference memory
	analogue and digital		medium	limited by analogue multipliers	permanent reference memory
	analogue		high		restricted by analogue reference decay

reduce the efficacy of analogue approaches, although analogue-analogue convolvers will always provide the maximum packing density in parallel implementations.

Table 1 summarises the features of the various approaches: performance figures are not included because of the difficulty of specifying hybrid systems and comparing performance parameters. Many trade-offs are evident from the Table but three particular cases appear most interesting:

(*a*) the parallel, digital convolver represents the best combination of speed and accuracy at the expense of power and size

(*b*) the parallel, analogue convolver represents the best combination of speed and packing density, at the expense of accuracy

(*c*) the serial, digital convolver[5] represents the best combination of packing density and accuracy for low-frequency applications

Of these three approaches, the parallel analogue convolver is the most promising for compact, lightweight applications requiring *T-B* products up to 1000.

3 Parallel convolver implementation

The remainder of the paper is devoted to a systematic study of the design philosophy regarding a particular realisation of an integrated, parallel, analogue convolver based on the principles of Section 2.2. Because a single-chip design only appears achievable using analogue-mode c.c.d./m.o.s., and because the convolver in this technology is an *active* component, then the design problem is very involved.[6,7] In this section, details of many of these considerations are outlined, pertaining to the convolver signal input, c.c.d. tapped delay line, the four-quadrant multipliers and a programmable reference memory. Although many other approaches to a single-chip convolver are possible, it is thought that the majority of the discussion will be relevant to other realisations based on c.c.d./m.o.s. technology.

3.1 MOS design techniques

In this Section, the requirements for the m.o.s. circuitry necessary to implement a convolver, when combined with a c.c.d. delay line, are outlined. Later work will present the limiting processes and appropriate equations that apply in the design of the integrated circuit.

M.O.S. linear circuitry concepts have not been fully exploited to date because of the problems associated with this approach. Hitherto, problems of noise, drift and low transconductance values (gain) have made their application to linear processing very restricted, in comparison to bipolar transistors. Although bipolar types have many useful operating parameter values, they do suffer from being relatively large in area. In the convolver design, where the object is to pack as many processing cells as possible on the chip, m.o.s. transistors would have the advantage of small size. Thus in the study leading up to the prototype convolver design, detailed later in the paper, the adverse features of m.o.s. structures were studied in their required configurations to see if any improvements could be innovated at the circuit level.

Fortunately, in the c.c.d. context, m.o.s. linear circuits can be made to appear to be high performance by adopting an elegant technique. The sampled-data nature of c.c.d. operation imposes on the m.o.s. circuitry its time quantisation. Thus the m.o.s. circuitry serving, say, the outputs of a convolver c.c.d. delay line, is required to be active for the time when the data is valid and then, temporarily, is not used. During these regular periods in which the data is being transferred in the c.c.d. (say about half of a clock period for a $2-\phi$ device) any drift in the non-sampled m.o.s. circuitry can be assessed and then eliminated by an appropriate correction. This could involve, for example, storing an error signal caused by transistor drift on a capacitor (store) so that it may be subtracted from a subsequent, new signal sample during the next valid data time. As the error correction process could be maintained at the c.c.d. clock p.r.f. the 'chopping' frequency could be very high and should be extremely effective in minimising any drift that would otherwise occur. It is also very useful as an approach because it could be used not only for correcting say the drift in a monolithic m.o.s. transistor summing amplifier, but also for reducing d.c. drift in the bias levels of the c.c.d. tap amplifiers and multipliers.

A further technique has been proposed[8] for improving the input/output linearity of c.c.d./m.o.s. circuits. Without this technique the gain and distortion in such circuits would be ill-defined, and lead to specification problems of the convolver system. The technique uses a single, high-gain, operational amplifier for improving the linearity by

including the otherwise nonlinear input circuit in the forward-gain loop of a voltage follower. Further details of this approach to linearising the input/output transfer function are given in the following Sections.

The basic operations of a parallel convolver for monolithic implementation have been discussed previously and are summarised in Fig. 1*b*. In order to achieve the necessary density of functions in an integrated circuit, all the operations such as delay, multiply and sum, must be obtained within a narrow strip on the surface of the silicon substrate. In practice, layout of the multipliers is a major problem. These should be fabricated at a pitch corresponding to the c.c.d. delay-line cell length, or in a multiple of it, for a neat device topology. In Section 3.4, considerations of an m.o.s. multiplier structure are detailed together with an error analysis for their performance.

In the remainder of this Section other details of the m.o.s. circuit design requirements are given, including a technique for storing and presenting the reference signal to the multipliers, and details of the summing circuitry. This work forms a unified basis for a fully integrated convolver design in c.c.d./m.o.s. technology.

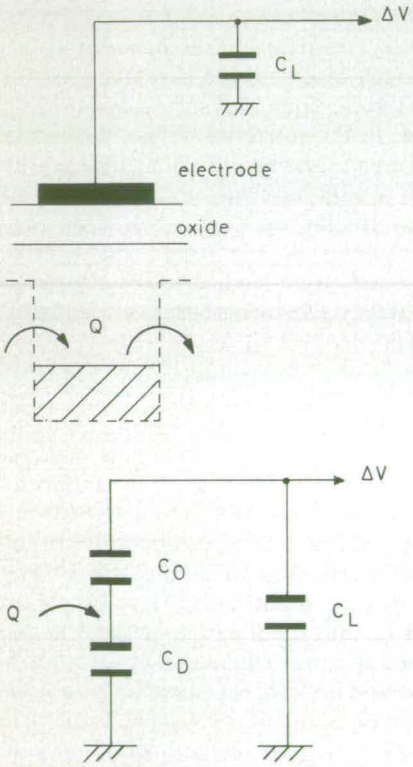


Fig. 2 *Principle of floating-gate reset charge sensing scheme*

$$\Delta V = \frac{QC_o}{C_oC_L + C_oC_L + C_oC_o}$$

$$\Delta V \approx \frac{Q}{C_L}$$

3.2 Signal register

A prime requirement of the transversal filter realisation of Fig. 1*b* is a periodically tapped, analogue delay line. Such sampled-signal delay functions are readily implemented using compact, low-power c.c.d. registers. These provide naturally the sequential delay of analogue signal samples, in the form of isolated charge packets, via a series of closely spaced controlling electrodes. The charge packets, which are amplitude modulated to contain signal information, are transferred along the register by driving the electrodes with m-phase clock waveforms (typically $m = 2, 3, 4$). Clearly, the time delay between samples is directly controlled by the clock frequency. An additional requirement of the tapped delay line, however, is that the information contained in each sample be periodically sensed whilst leaving the charge packet isolated and free to transfer along the register.

A non-destructive sensing scheme[9] which may be implemented in any c.c.d. technology is illustrated in Fig. 2. A conventional clocked electrode in the c.c.d. structure is replaced by a 'floating-gate' reset (f.g.r.) sense electrode. This electrode is firstly reset to a bias voltage whilst there is no signal charge underneath it. It is then isolated and any charge subsequently injected (as a result of the c.c.d. clocking action) causes a related change in electrode potential. The signal may be sensed via an m.o.s. transistor amplifier to provide a voltage output at a low impedance.

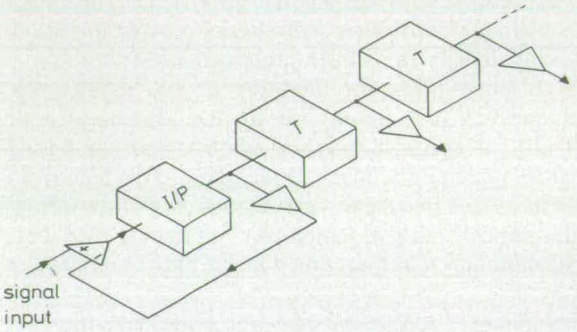


Fig. 3 *Feedback linearisation of the c.c.d. transfer function*

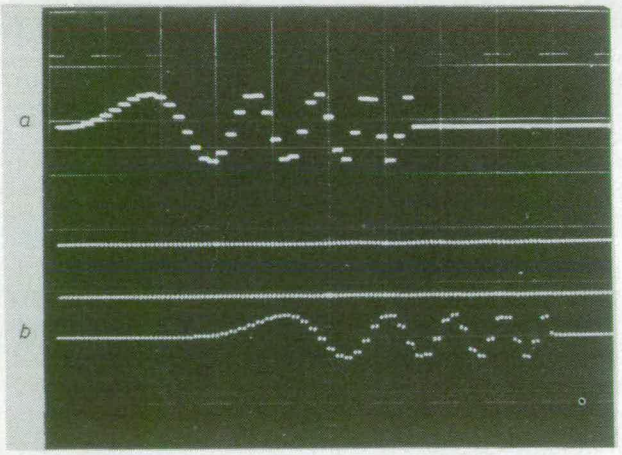


Fig. 4 *FGR tap output showing a delayed chirp waveform*

a Sampled input waveform
b Delayed tap output (2 V, 200 μs)

Although the intrinsic transfer function may be complex the linearity and gain of this structure may be well defined by applying a feedback linearisation scheme[8] at the register input, as shown in Fig. 3. In this case an f.g.r. tap is used to meter the input charge level which is then corrected via a differential amplifier until the desired tap output signal is obtained. The charge packet is isolated and transferred along the c.c.d. delay line resulting in identical output signals at all subsequent taps, which are replicas of the input metering tap.

A typical delayed tap output resulting from a chirp input signal is shown in Fig. 4. Note that the output signal is 'chopped' and the true signal information appears for only a fraction of the complete clock period due to the reset-isolate-inject charge sensing sequence.

Considering errors within the structure, the c.c.d. tapped delay element has three fundamental disadvantages:

(i) *Transfer inefficiency.* A small fraction of each charge packet is left behind at each transfer, causing cumulative signal degradation along the register. For an inefficiency ϵ per tapped stage, then the output at a non-destructive tap n at time m is

$$\sum_{r=0}^{r=(m-n)} \frac{(n+r)!}{n!r!} \epsilon^r (1-\epsilon)^n s_{m-n-r}$$

In the frequency domain this has the effect of imposing a lowpass filter characteristic, the severity of which increases with increasing cumulative transfer inefficiency.

(ii) *Dark current.* As signal charge packets travel down the c.c.d. they are subject to an additive error due to minority carriers,[10] which are thermally generated within the semiconductor. For a uniformly clocked c.c.d. the net effect is a steady increase in signal charge with time. As the effect is essentially linear there results a constant offset error at tap n of

$$\hat{V}_{sig} \frac{J_D A_C T_t n}{\hat{Q}_{sig}} \text{ volts}$$

where $\hat{V}_{sig}$ is the change in tap output potential corresponding to the peak charge capacity $\hat{Q}_{sig}$, J_D is the dark current density, A_C is the average c.c.d. cell area and T_t is the time delay between taps.

(iii) *Amplifier mismatching.* Although the c.c.d. transfer function can be linearised at the input prior to launching the signal charge packet, the ultimate accuracy of the tapped delay line requires that the characteristics of subsequent tap amplifiers match those of the input metering tap. Such mismatches may be caused by random gain and threshold variations in the m.o.s. transistors. It is possible, however, to design the buffer amplifiers such that these effects result only in quiescent offset errors. It is shown later that these errors may be cancelled at the output.

3.3 Reference register

As the c.c.d. signal register provides the necessary time-shift process a static reference register, permanently driving one set of multiplier ports, is sufficient. The most compact analogue reference store consists of a series of capacitors followed by buffer amplifiers, if necessary, to drive the multiplier circuitry. Many reference elements may be loaded from a common reference input busbar using on-chip digital multiplex techniques, or by reading the reference serially into a c.c.d. register and then transferring the reference set in parallel to the capacitor array. The former technique, illustrated in Fig. 5, is preferred for two reasons:

(*a*) Real-time updating permits feedback linearisation of each reference sample to be implemented right up to the appropriate multiplier port.

(*b*) By making the multiplex circuitry addressable individual reference addresses may be selected randomly as well as sequentially, which may be of great use in systems requiring permuted data (prime transforms, for example).

The major limitation of the capacitively-held reference is information decay due to charge leakage. The dominant contribution toward leakage current occurs at the diffusion terminal of the sampling transistor, and its magnitude depends upon the characteristics of the diffused diode in a complex manner. Assuming, for simplicity, a constant leakage current I_L then the stored value r decays such that

$$r \rightarrow r - \frac{I_L}{C}(t - t_r) \tag{2}$$

where C is the holding capacitance and t_r is the last refresh time.

Capacitive breakthrough of the sample waveform onto the hold capacitor via gate overlap capacitance is a second-order effect. The voltage level at which the sample transistor cuts off, and thus the magnitude of the breakthrough, is a function of the voltage stored. Assuming a linear buffer amplifier characteristic, the net effect is a quiescent offset combined with an apparent reduction in gain

$$r \rightarrow (V_0 - r_0 - V_T)\frac{C_f A_v}{C_f + C_h} + r\left(1 - \frac{C_f A_v}{C_f + C_h}\right) \tag{3}$$

where r_0 is the reference zero level, V_0 is the off potential of the sampling waveform, V_T is the threshold voltage of the sampling transistor, A_v is the small-signal voltage gain of the buffer amplifier and C_f and C_h are the feedthrough and hold capacitances, respectively.

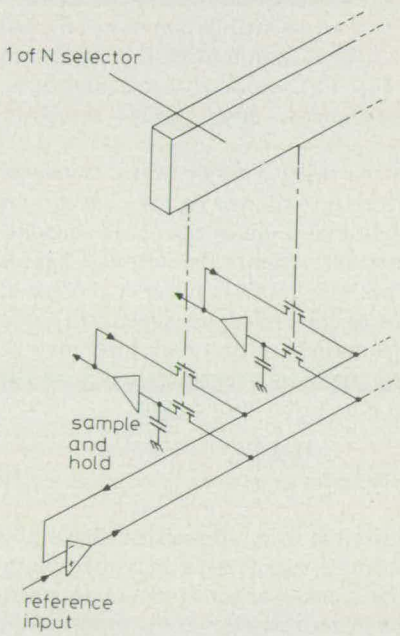


Fig. 5 *Digitally multiplexed capacitive analogue reference register with feedback linearisation*

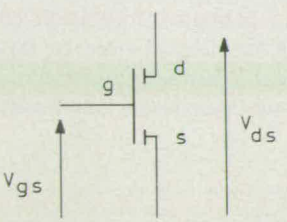


Fig. 6 *M.O.S. transistor nomenclature*

3.4 Multipliers

For general-purpose signal processing four-quadrant multiplication of the signal and reference voltages is required at each convolution point. An economical multiplication technique is based upon the essentially linear transconductance of an m.o.s. transistor operating in the non-saturation region. A first-order expression for the drain current of such a transistor, shown in Fig. 6, is given by

$$I_D = \beta_m \left\{ (V_{GS} - V_T) V_{DS} - \frac{V_{DS}^2}{2} \right\} \tag{4}$$

from which the change in drain current obtained upon disturbing the gate voltage by an amount V'_{GS} is

$$\Delta I_D = \beta_m V_{DS} V'_{GS} \tag{5}$$

It is worthwhile noting that V'_{GS} may take either sign and that eqn. 5 is equally valid if V_{DS}, as defined in Fig. 6, changes sign and the drain and source terminals become transposed.* Thus the desired four-quadrant multiplication may be obtained by applying the operand voltages

(i) as a constant v_r with respect to some fixed 'zero' level V_{r0} at one diffused terminal

(ii) as a change in gate voltage v_s with respect to some 'zero' level V_{s0}, and by sensing the corresponding change in drain current.

Previously the difference voltage v_s has been synthesised[6,7] by using two identical transistors with a common diffused terminal (v_r); the gate of the second transistor being driven with the 'zero' voltage V_{s0}. The accuracy of this method is, however, critically dependent upon good transistor matching.

An alternative approach using a single m.o.s. transistor is feasible where the reference voltage (v_r) is constant, as in this case. Here the undesired terms in eqn. 4 also become static, representing a constant offset at the output. Changes in the gate (signal) voltage now result in product changes at the output with respect to the (reference-dependent) zero level. Where necessary a defined output zero level may be obtained by alternatively applying a signal zero V_{s0} to the transistor gate to obtain the output zero circuit:

$$I_0 = \beta_m \left\{ (V_{s0} - V_T) V_r - \frac{V_r^2}{2} \right\} \tag{6}$$

and then subtracting this value from subsequent 'signal plus zero' samples. The alternate signal zero may be simply obtained by tapping the c.c.d. register at alternate stages and multiplexing signal and zero samples at the input.

Thus a form of chopper stabilisation may be applied at the output to set and hold a zero level during the reset period, and output, subsequently, the product signal as a change in this level. Not only does this zero multiplexing technique remove the transistor matching requirement (a time-multiplexed transistor is perfectly matched to itself) but the alternate zero levels also contain information on all quiescent offset errors in the signal port (due to tap-amplifier mismatching and threshold voltage variations in the multiplier transistor), which are thus also automatically removed at the output.

* Because the drain and source terminals remain at a constant potential there is no additional modulation of the drain current due to substrate bias effects so that eqn. 5 remains valid for more complex drain current expressions than eqn. 4.

Two possible errors arise when using this transconductance multiplication technique

(*a*) The changes in drain current may modulate the reference voltage v_r due to the finite output admittance of the reference amplifiers. If the reference value has been set (using a feedback linearisation scheme) during the signal zero phase then the product error due to a reference amplifier output admittance of g_0 is given by

$$sr \rightarrow sr - s^2 r \frac{\beta_m}{g_0} \tag{7}$$

where the error factor β_m/g_0 may be typically 1%.

(*b*) Individual 'matched' transistor gain factors β_m may vary within a given integrated circuit by as much as 5% and between different circuits on the same slice by 20%. The variation between devices may be trimmed out at individual summing amplifiers, but there still remains a random product gain variation $\delta\beta/\beta$ between convolution points on any given device. The statistics of this variation are obviously process dependent but the effect may be minimised by using multiplication transistors of large area and may be cancelled out in an adaptive system.

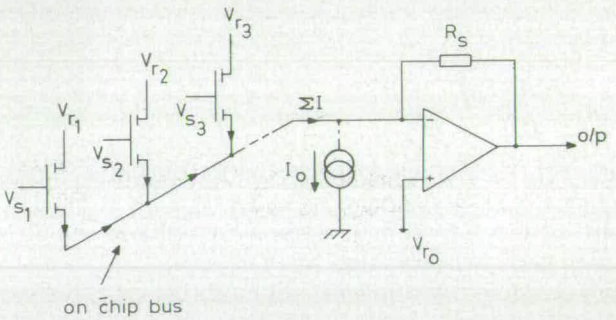


Fig. 7 *Multiplier summing arrangement*

$V_s = v_s + V_{s0}$

$V_r = v_r + V_{r0}$

3.5 Summing

The provision of the multiplier reference voltage V_{r0} in addition to the summation of the product currents may be conveniently realised using a single operational amplifier and feedback summing resistor as shown in Fig. 7. Note the single current summing bus which again allows a compact circuit topology.

The removal of the alternate zero current level may be easily accomplished at this stage by programming a current source/sink during alternate periods to remove the zero current level from the summing busbar. The removal of this quiescent current before the summing resistor enables the wanted output signal to occupy the full dynamic range of the summing amplifier rather than a restricted portion of it, thus enhancing the noise figure of the complete device and allowing the output gain to be varied without a corresponding d.c. shift.

Generally, the summing amplifier is a limiting factor in terms of system speed in that its output must settle within the time period for which the signal is true; for an m-phase device the settling time must therefore be $T_c s/m$ where T_c is the c.c.d. clock period. If a signal sample-hold facility is made available on-chip prior to the multiplier gate, this settling time may be extended. Although the intrinsic

c.c.d./m.o.s. circuitry may be designed to operate at clock frequencies up to 10 MHz, it is difficult to achieve sampling rates at the summing amplifier much above 1 MHz without resorting to bulky, high-power amplifiers which make mobile, lightweight applications less attractive.

An operational amplifier design for this purpose in m.o.s. technology has been reported[11] using single-channel transistors. It was fabricated with an N-channel c.c.d. process and gave satisfactory performance to above 1 MHz. A chopper stabilisation technique, similar to that outlined in Section 3.1, gave a useful drift figure of < 1 mV over an 80°C change in ambient temperature. The area of the amplifier was only 0·6 mm square, making it suitable for many c.c.d. peripheral functions.

3.6 *Integrated error effects*

Indications have been given so far, in this section, of the major contributions towards inaccuracy in this particular realisation of the convolution sum. The effects of these error sources at the output of the convolver are now discussed with reference to other work.[12] Unfortunately, unless precise information about both the signal and reference waveforms is available quantitative analysis is not feasible; it is, however, useful to review the general effects of these errors so that potential problem areas may be analysed more fully.

Charge transfer inefficiency (ϵ) in the c.c.d. register has already been shown to degrade signal information at high frequencies; this being a cumulative effect with increasing $N\epsilon$ products. In terms of frequency filtering applications, however, the net effect at low frequencies is a shift in the transition edges between pass-band and stop-band of a factor $(1-\epsilon)$, regardless of the number of filter stages.[13] The effect upon correlator performance is somewhat more signal dependent, but charge transfer inefficiency is known to cause a slump in the correlation peak and a corresponding increase in sidelobe significance. It has been suggested that in certain correlation applications an $N\epsilon$ product of 2 is tolerable.[14]

Random gain errors in the multiplication process are analogous to tap weight errors in split-gate, c.c.d. filters and as such may be expected to impair stop-band suppression in frequency filtering applications. Correlation applications are more tolerant to these errors which become attenuated by the 'processing gain' ($N^{1/2}$) of the filter.

Quiescent offset errors in both the signal and reference channels appear, as such, at the convolver output for suitably large time-bandwidth products. Errors of this nature in the signal channel are automatically removed with the multiplication technique outlined in Section 3.4. Quiescent reference errors Δr produce a net output error $\bar{s}\overline{\Delta r}$ where the bar denotes an average value. Clearly, such errors may be reduced or eliminated where either sequence can be chosen to have zero mean value.

The signal distortion term imposed by the finite output admittance of the multiplier driving stages unfortunately precludes a general linear analysis of the resultant error. The effect of this distortion may, however, be demonstrated in frequency filtering applications where a pure sinusoid is applied at the input. Then, from eqn. 7, we may expect at the output d.c. (zero Hz) and second harmonic distortion components of magnitude $\frac{1}{2}\beta_m |H(0)|/g_0$ and $\frac{1}{2}\beta_m |H(2\omega)|g_0$ respectively. In correlator applications additional correlation peaks are to be expected whenever a match is detected between the generated harmonic signal component and the reference waveform.

The refreshing and subsequent decay of reference values results in a modulation of the output waveform. As this process is not generally synchronised with the incoming data sequences it appears as a form of noise at the output. Naturally, the magnitude of this noise is directly related to the decay rate and update frequency of the reference samples; these factors determine the maximum number of reference points which may be updated sequentially. Thereafter reference refreshing must take place in parallel blocks.

Although several potential error sources have been identified here it is clear that not all of them will be applicable to any one application of the device. In general, frequency filtering applications requiring large stop-band attenuations are more sensitive to these errors. Matched filtering (correlation) applications, however, are considerably more tolerant to random errors which become attenuated by the 'processing gain' of the filter; generally, the only significant errors are those which correlate with either the signal or reference sequences.

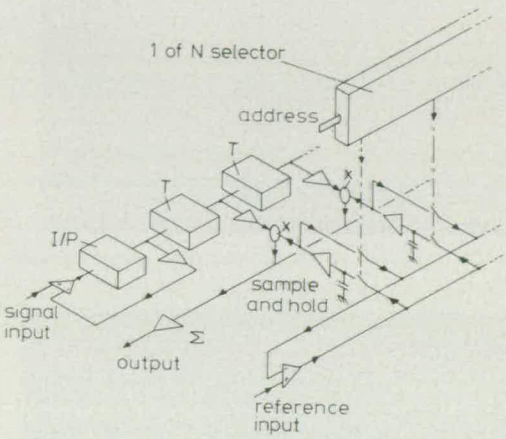


Fig. 8 *Block diagram of a programmable transversal filter*

4 Implementation

The parallel system circuit principles outlined in this paper have been applied to the design of a monolithic programmable transversal filter for use in a variety of signal processing systems. A block diagram of the device structure is shown in Fig. 8, which may be compared directly with the system diagram of Fig. 1*b*. The tapped c.c.d. delay line feeds one port of a register of m.o.s.t. multiplier elements. The reference port is similarly driven from the static register, the elements of which are updated individually from a multiplexed analogue reference input busbar. Provision is made for the reference values to be feedback linearised at the multiplier ports via a similarly multiplexed feedback busbar.

The feasibility of this filter structure has previously been verified by construction of 32-point, hybrid correlator[15] using a discrete c.c.d. tapped delay line and monolithic m.o.s.t. multipliers. The design of a completely monolithic 64-point, programmable transversal filter, shown in Fig. 9 has been completed. Initial results confirm operation of the device as a fully programmable transversal filter over sampling frequencies in the range 500 Hz to 100 kHz. Fig. 10 demonstrates the autocorrelation of a chirp waveform using the filter. The output waveform

agrees well with the predicted sinc x response, verifying the accuracy of the multipliers, and demonstrates a dynamic range of 50 dB. The upper trace shows the output from the last c.c.d. stage which appears very suitable for cascading second and subsequent devices in 64-point blocks. The complete integrated circuit measures 4·5 mm × 3·3 mm, demonstrating the high packing density achievable with this design technique.

An updated design is planned on a two-level, polysilicon process with a potentially higher packing density which is expected to yield a 128-point correlator on a chip less than 5 mm square. These devices are cascadable, and as such a single-board correlator of up to 1000 points may be envisaged. However, before definitive predictions of filter performance can be made, considerable research and characterisation work must be undertaken, based on the prototype results.

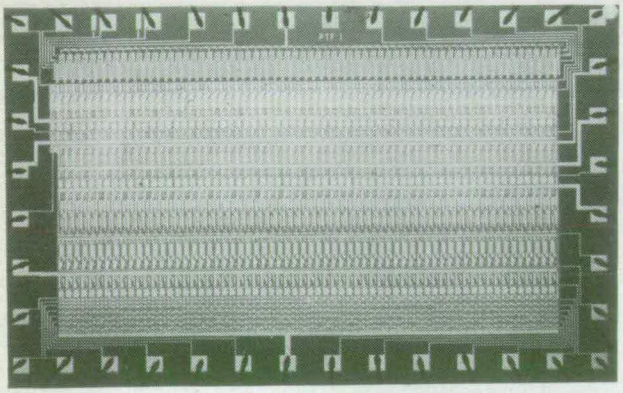


Fig. 9 *Monolithic 64-point analogue-analogue programmable transversal filter*

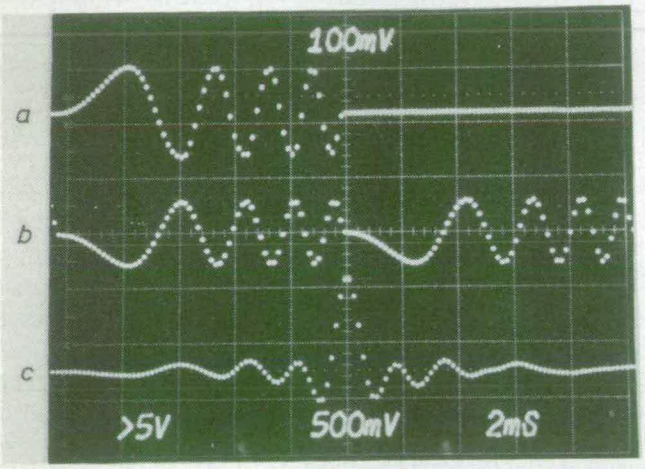


Fig. 10 *Matched filtering of chirp waveform using the 64-point filter of Fig. 9*

a 64th tap output
b reference waveform
c filter output

5 Conclusions

It is clear from the ever advancing progress of integrated circuit development that developments in the hardware implementation of systems will lead to more compact and effective realisations. In this paper, convolver modules have been developed which exploit the potential of custom c.c.d./m.o.s. technology in analogue form. Although developments according to this approach look of great potential presently, undoubtedly the availability of compact, high performance digital multiplier hardware would seriously affect the viability of the analogue approach. However, where space and weight is at a premium, the realisation of a convolver with modest performance may still be a useful proposition. Clearly, the future of the analogue approach is dependent upon the market requirements and the interest shown in compact processors.

This study has indicated that analogue device design has many possibilities when directed to process time-sampled data. It has further shown that an analysis of the errors in such a system is imperative before the operation of an analogue convolver can be properly assessed. Such work needs to be extended in the light of other prototype measurements before the full limitations of the approach can be determined.

6 Acknowledgments

This work has been sponsored at various times by several bodies including AUWE, DCVD, Procurement Executive, UK Ministry of Defence and the UK Science Research Council. The authors are grateful to the following for their contributions: D.J. MacLennan, G.F. Vanstone, J.W. Arthur, M.A. Jack and D. Saxton. The facilities of the Wolfson Microelectronics Liaison Unit, at the University of Edinburgh, are gratefully acknowledged.

7 References

1 BOYLE, W.S., and SMITH, G.E.: 'Charge-coupled semiconductor devices', *Bell Syst. Tech. J.*, 1970, **49**, pp. 587–593
2 COPELAND, M.A.: 'Interaction between microprocessors and custom LSI' *in* 'Microprocessors and their applications' AGARD Lecture Series 87, New York 1977
3 SQUIRE, W.D., WHITEHOUSE, J.H., and ALSUP, J.M.: 'Linear signal processing and ultrasonic transversal filters', *IEEE Trans.* 1969, **MTT-17**, pp. 1020–1040
4 HERRMANN, E.P., GANDOLFO, D.A., BOORNARD, A., and STEPPS, D.B.: 'CCD programmable correlator', Proceedings of the international conference on the technique and applications of c.c.d.s, Edinburgh, 1976, pp. 232–237
5 MAVOR, J., JACK, M.A., SAXTON, D., and GRANT, P.M.: 'Design and performance of a programmable real-time charge-coupled device recirculating delay-line correlator', *IEE J. Electron. Circuits & Syst.*, 1977, **1**, pp. 137–144
6 HARP, J.G., VANSTONE, G.F., MACLENNAN, D.J., and MAVOR, J.: 'Analogue correlators using charge-coupled devices', Proceedings of the c.c.d. applications conference, San Diego, 1975, pp. 229–235
7 BOSSHART, P.: 'An integrated analogue correlator using charge-coupled devices', *IEEE ISSCC*, 1976, **76**, pp. 198–199
8 MACLENNAN, D.J., and MAVOR, J.: 'Novel technique for the linearisation of charge-coupled devices', *Electron. Lett.*, 1975, **11**, pp. 222–223
9 DENYER, P.B., and MAVOR, J.: 'Design of CCD delay lines with floating-gate taps', *IEE J. Solid-State & Electron Devices*, 1977, **1**, pp. 121–129
10 TASCH, A.F., BRODERSEN, R.W., BUSS, D.D., and BATE, R.T.: 'Dark current and storage-time considerations in charge-coupled devices', Proceedings of the c.c.d. applications conference, San Diego, 1973, pp. 179–188
11 WESTE, N., and MAVOR, J.: 'MOST amplifiers for performing peripheral integrated circuit functions', *IEE J. Electron. Circuits & Syst.*, 1977, **1**, pp. 165–172
12 MACLENNAN, D.J., MAVOR, J., and YEOW, Y.T.: 'Errors in programmable CCD transversal filters and correlators', Proceedings of the international conference on the technique and applications of c.c.d.s, Edinburgh, 1976, pp. 259–268
13 BUSS, D.D., BAILEY, W.H., HOLMES, J.D., and HITE, L.R.: 'Charge transfer device transversal filters for communication systems', *Microelectronics*, 1975, **7**, pp. 46–53
14 BUSS, D.D., COLLINS, D.R. BAILEY, W.H., and REEVES, C.R.: 'Transversal filtering using charge transfer devices', *IEEE Trans.*, 1973, **SC-8**, pp. 138–146
15 MAVOR, J., ARTHUR, J.W., and DENYER, P.B.: 'Analogue CCD correlator using monolithic MOST multipliers', *Electron. Lett.*, 1977, **13**, pp. 373–374

MONOLITHIC, PROGRAMMABLE ANALOGUE C.C.D. TRANSVERSAL FILTER

Indexing terms: Charge-coupled-device circuits, Correlators, Filters

A 64-point monolithic, fully programmable analogue c.c.d. transversal filter has been produced. The design principles and limitations of this approach are discussed, and results are presented for the filter acting as a chirp-waveform correlator and for other filtering functions.

Introduction and description: The role of charge-coupled devices as tapped analogue delay lines for fixed-impulse-response transversal filters is well established. However, realisations of variable-response filters—or programmable forms—are still in early development, owing to the complexity of the peripheral circuitry necessary to perform the variable weighting of signal samples. Digital weighting coefficients have been realised and reported.[1] However, they occupy significantly large areas for anything other than a binary reference. Analogue weighting coefficients are more desirable for general signal processing and potentially offer optimum packing density for minimum package counts. Realisation of analogue weighting implies 4-quadrant multiplication, in addition to an analogue reference store,[2] and, until now, the difficulties of realising these functions have limited the development of prototype, monolithic devices.[6]

This letter reports initial results for an operating c.c.d/m.o.s.t., analogue, programmable transversal filter in integrated circuit form. The principles of this approach have been reported previously in hybrid form.[3] A block diagram summarising the main features of monolithic circuit is given in Fig. 1. A 128-stage, dynamic 3-phase c.c.d. signal register is used, which is tapped at alternate stages to provide 64 filter points. A floating-gate reset (f.g.r.) technique[4] is used for implementing the multitapped signal delay line. Provision is made for feedback linearisation[5] of the c.c.d. transfer function by using an offchip operational amplifier.

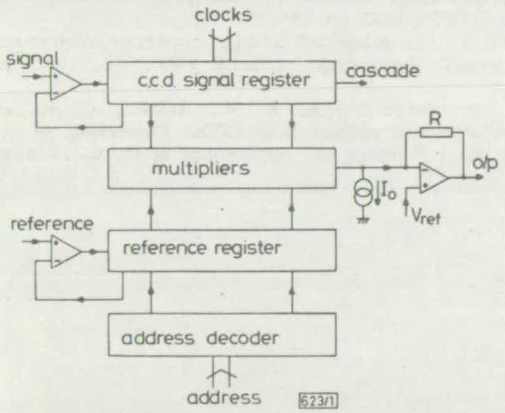


Fig. 1 *Block diagram of programmable transversal filter*

The reference weights are supplied in parallel form from a static onchip analogue memory, realised by using a register of m.o.s. capacitors and buffer amplifiers. The reference memory may be updated serially (or in any desired order) via a digitally multiplexed analogue input line; again, provision is made for linearisation. Onchip logic decodes an external address busbar to enable the required reference sample to be refreshed.

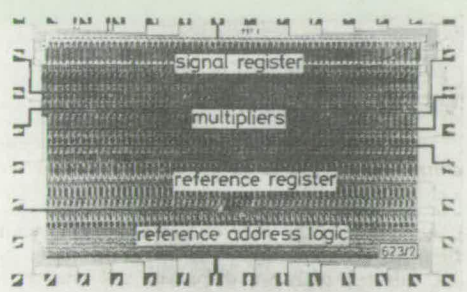


Fig. 2 *Monolithic 64-point programmable transversal filter*

The reference and signal samples are multiplied by using a single m.o.s. transistor at each point, which operates in the linear transconductance region. True 4-quadrant analogue multiplication is obtained by using a novel time-multiplexing technique which gives improved accuracy and dynamic range over conventional matched-transistor realisations.[2,6]

A photograph of the completed chip is given in Fig. 2, which shows that an extremely efficient use of silicon area can be achieved with this approach; the 64-point filter occupies an area of 4·7 × 3·2 mm.

Multiplication: Accurate 4-quadrant multiplication is attained at each filter point by using single m.o.s. transistors, according to the following theory. For an m.o.s. transistor operating in the presaturation region, we may write

$$I_{DS} = \beta_m [(V_{GS} - V_T)V_{DS} - V_{DS}^2/2] \quad (1)$$

where β_m is a process-dependent constant and the other symbols have their usual meanings. For a given (constant) drain-source voltage V_{DS}, any change in gate voltage V_{GS} stimulates a change in drain current:

$$\Delta I_{DS} = \beta_m V_y V_x \quad (2)$$

where $V_y = V_{DS}$ and $V_x = \Delta V_{GS}$, for all sign combinations.

Four-quadrant multiplication is thus obtained at each filter point by applying the reference voltage at the transistor drain, with respect to a constant-voltage source, and by applying the signal voltage as a change in potential at the transistor gate, with respect to some quiescent 'zero' value. All currents are summed on a single common-source busbar, which is held at the required constant potential by the output summing amplifier (offchip). A true output level is obtained by time multiplexing the multiplier transistors to receive alternately 'zero' and 'signal-plus-zero' gate voltages, and by programming a current source at the output to detect and subtract the unwanted current component I_0 during the zero phase. During the subsequent signal-plus-zero phase the resultant output is simply the sum of products term required. The time-multiplexed zero arrangement is conveniently realised by exploiting the alternately-tapped c.c.d. register via the application of alternate zeros at its input. The parallel tap outputs then alternate between zero and signal plus zero, so that the filter output is constantly corrected at half the c.c.d. clock frequency.

This time-multiplexed multiplication scheme offers three practical advantages over conventional matched-transistor-pair techniques:[2,6]

(*a*) Circuit layout is reduced and simplified.

(*b*) The need for matched transistors is removed, resulting in greater multiplication accuracy. Also, all quiescent offset errors in the signal channel are cancelled.

(*c*) The facility for leaking off the unwanted current component, before the summing amplifier, allows use of the full dynamic range of this element.

Results: Initial results confirm operation of the device as a fully programmable transversal filter over sampling frequencies in the range 500 Hz–100 kHz or greater. Fig. 3 demonstrates the filter used for the autocorrelation of a chirp waveform. The output waveform agrees well with that predicted mathematically, verifying the accuracy of the new multiplication technique. The upper trace shows the output from the last (128th) c.c.d. stage, which appears very suitable for cascading second and subsequent devices in 64-point blocks.

The volatility of onchip analogue memory implies that the reference samples require refreshing periodically. This was performed continuously in the example, in synchronisation with the c.c.d. clock (although convenient, this is not a necessary restriction). In subsequent experiments, correlation peaks have been maintained in uncorrupted form for reference decay times in excess of 5 s. Initial measurements suggest a dynamic range of 60 dB at the output for the correlation of two square waves (maximum theoretical output). For

sinusoids this value is reduced by 6 dB, owing to the lower r.m.s. value of these waveforms. The accuracy and dynamic range of the device are such that a subjectively useful impulse response is obtained by launching an impulse down the c.c.d. register; the output is then a serial read-out of the contents of the reference register.

The device is not restricted to correlator or matched-filter applications. References may be programmed to realise frequency-filtering functions by simply applying the inverse Fourier transform of the desired frequency characteristic as the reference sequence, i.e. impulse response. Operation of the device in this mode has been confirmed by programming a basic sinc X reference function to realise a lowpass-

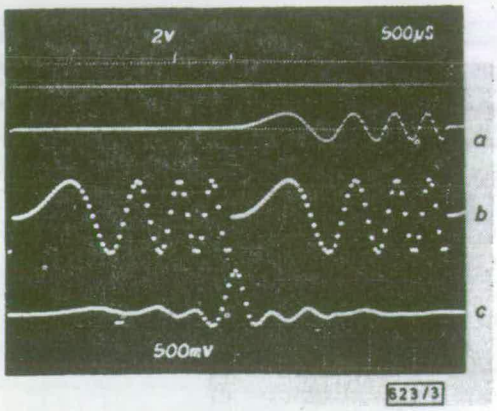


Fig. 3 *Matched filtering of chirp waveform*

Chirp 0–$f_s/8$ hertz; f_s = 40 kHz
(*a*) 64th-tap c.c.d output, (*b*) Reference waveform,
(*c*) Correlator output

filter characteristic. These applications are known to be more sensitive to multiplier accuracy and are a good indication of the quality of the device.

A full characterisation programme is under way to determine the operational limits of the device; results will be reported later.

Conclusions: The fully programmable analogue transversal filter is a promising realisation of a powerful general-purpose signal-processing element. A monolithic 64-point prototype device using a novel multiplication technique has been produced. Initial results confirm operation of the device in both frequency- and matched-filtering configurations over a range of frequencies suitable for sonar and other submegahertz applications. Other potential applications include spectrum analysis using the chirp-*Z* or prime-transform techniques.[7]

Acknowledgments: This work was sponsored by AUWE(S), Portland. The authors are grateful to the Plessey Company Ltd., for processing the devices, and to the Wolfson Microelectronics Liaison Unit, Edinburgh, for design and test facilities.

P. B. DENYER
J. W. ARTHUR
Wolfson Microelectronics Liaison Unit
University of Edinburgh
Mayfield Road, Edinburgh EH9 3JL, Scotland

J. MAVOR
Department of Electronic & Electrical Engineering
University of Edinburgh
Mayfield Road, Edinburgh EH9 3JL, Scotland

20th October 1977

References

1 COPELAND, M. A. : 'Interaction between microprocessors and custom LSI'. AGARD Lecture Series, 1977, 87, New York
2 HARP, J. G., VANSTONE, G. F., MACLENNAN, D. J., and MAVOR, J.: 'Analogue correlators using charge coupled devices'. Proceedings of conference on the application of CCDs, 1975, pp. 229–235
3 MAVOR, J., ARTHUR, J. W., and DENYER, P. B.: 'Analogue c.c.d. correlator using monolithic m.o.s. multipliers', *Electron. Lett.*, 1977, 13, pp. 373–374
4 DENYER, P. B., and MAVOR, J.: 'Design of CCD delay lines with floating-gate taps', *IEE J. Solid-State & Electron Devices*, 1977, 1, pp. 121–129
5 MACLENNAN, D. J., and MAVOR, J.: 'Linearisation of the charge coupled device transfer function'. Proceedings of conference on the application of CCDs, 1975, pp. 291–294
6 BOSSHART, P.: 'An integrated analogue correlator using charge-coupled devices'. Proceedings of IEEE ISSCC 76, 1976, pp. 198–199
7 BUSS, D. D., BRODERSEN, R. W., HEWES, C. R., and DEWIT, M.: 'Spectral analysis using CCDs'. Proceedings of conference on the technology and applications of CCDs, 1976, pp. 208–218

NOVEL M.O.S. DIFFERENTIAL AMPLIFIER FOR SAMPLED-DATA APPLICATIONS

Indexing terms: Differential amplifiers, Field-effect integrated circuits, Linear integrated circuits

A novel form of differential amplifier is reported which employs a simple, capacitive input differential circuit in place of the conventional differential pair and current source. The amplifier is especially suitable for monolithic integration in an m.o.s. technology, and generates a time-multiplexed, sampled-data output signal compatible with many current signal-processing techniques.

Introduction: The requirement for compact, low-power differential amplifiers and comparators for use in m.o.s.-based integrated circuits is well recognised. These functions are most often required in monolithic implementations of data convertors and in the general class of signal processors realisable with charge-coupled and bucket-brigade devices. The aim is clear: to produce a stable amplifier with repeatable characteristics capable of operating over bandwidths into the low megahertz range. The limitations, however, are formidable: large variations in transistor gain and threshold voltage make stable biasing and operation difficult to achieve in practice.

M.O.S. operational amplifiers reported to date[1-3] have been based on conventional differential circuits requiring matched transistor pairs and constant-current sources; the ultimate performance of the amplifier depends critically on the quality of these components.

The amplifier reported here is based on a new differential stage realisation employing a simple, capacitive potential divider which is combined with switching transistors to provide a true difference signal. The differential circuit is ideally suited to implementation in any m.o.s. technology, is insensitive to bias and common-mode signals and is automatically stabilised for drift-free operation.

Principle: The new differential technique, illustrated in Fig. 1, is based on the capacitive potential divider. This will respond to a change in potential at node A with a corresponding change in potential at node B of

$$\Delta V_B = \frac{C_1}{C_1 + C_2} \Delta V_A \qquad (1)$$

regardless of the quiescent potentials already existing at these nodes. The operation of the circuit in a differential mode is as follows:

(*a*) With the clock ϕ 'on', node A is connected to one differential input V_+ and simultaneously B is reset to the required 'zero' level V_R.

(*b*) As ϕ turns 'off', node B is isolated and, at this instant, the V_+ value is effectively sampled and remembered; any subsequent change in potential with respect to this value will result in a related change in potential at node B, via eqn. 1.

(*c*) Node A is then connected to the V_- terminal as the complementary clock $\bar{\phi}$ turns 'on', causing a resultant change in the potential at node A of $V_- - V_+$. As the potential at the isolated node B must track this signal, the true difference in potential between the input terminals has been found, whereas the common-mode signal is rejected during the reset phase.

The overall cycle thus consists of a reset phase, during which a 'zero' signal appears at the output, followed by a signal phase, during which the difference in potential between the two input terminals appears at the output. The chopped nature of the output signal is not disadvantageous in any of the applications envisaged, which all operate in a sampled-data mode. Indeed, such operation is desirable in, for example, many c.c.d. and data convertor systems. In systems employing feedback, the closed loop is free to settle continuously during the second phase of the clock period, making possible a true sampled-data operational amplifier.

Although the gain factor in eqn. 1 must always be less than unity, the differential stage is readily followed by high-impedance single-input gain stages and subsequent output stages as required. Cascaded m.o.s.t. inverter stages having gains of up to 50 dB are quite feasible and have been demonstrated previously.[2]

Drift, bias and offset: The problem of drift in the differential stage is completely eliminated by the reset action at the output node. There is also no error associated with this stage because both input signals are time multiplexed onto the same potential divider, which is inherently matched to itself. The following m.o.s. amplification stages will, however, suffer from bias and drift problems, which cause shifts in the output zero level and thus apparent offsets at the input. Self-tracking bias circuits and chopper stabilisation techniques have been suggested elsewhere,[1-3] but a simple technique which appears most suitable in this case is indicated in Fig. 1, by the connection shown as a broken line.

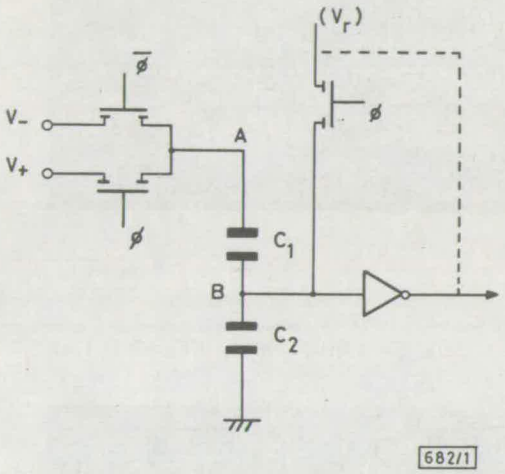


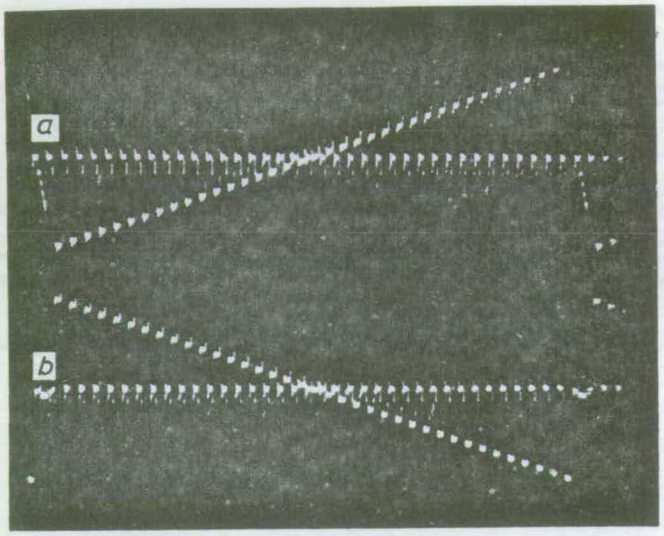
Fig. 1 *Principle of differential amplifier*

Instead of resetting to a bias voltage V_R, the inverting amplification stage is merely short circuited. Suitable design of the inverter transfer function can allow the $V_{OUT} = V_{IN}$ condition to reset the amplifier in the required operating range. Threshold voltage variations of, say, ±300 mV, will cause corresponding changes in the output bias point, but these represent referred offsets of as little as ±1 mV at the input for an open loop gain of 50 dB. The need for a separately generated bias voltage is thus removed and the technique is again notably simple to implement.

Speed: The bandwidth of the system is limited (by Shannon's sampling theorem) to one-half of the clock sampling frequency. The charging time at the capacitive input or the settling time of any feedback loop then decides the minimum time requirement in each clock phase. The loop settling time will generally be greater; closed-loop settling times of 1 to 2 μs have been reported for m.o.s. amplifiers driving 70 pF loads.[3] For smaller onchip loads of 10 pF or less, we may expect target settling times of the order of 250 ns, implying the use of a 2 MHz biphase clock and realising a signal bandwidth of 1 MHz.

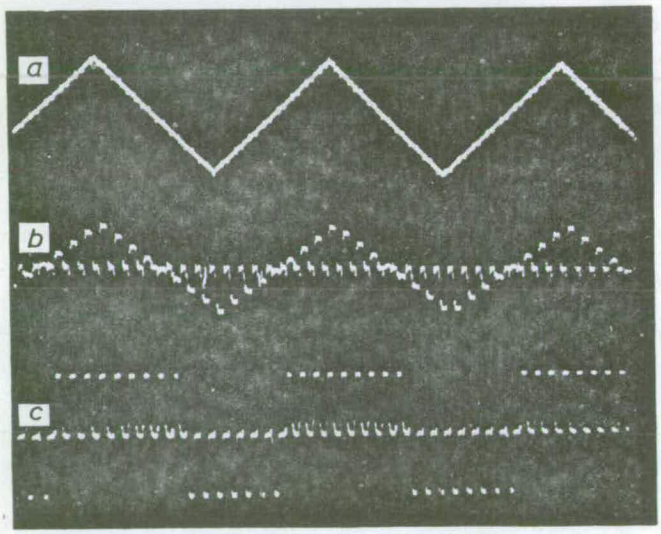
Experimental results: To verify the principles outlined above a differential stage of the type shown in Fig. 1 has been constructed by using discrete components. Complementary clock waveforms of amplitude 15 V were generated to drive the stage at 2 MHz. Fig. 2*a* shows the output of the differential stage in response to a ramp input at one input terminal with a d.c. reference at the other. Fig. 2*b* shows the equivalent output of the stage with the input terminals exchanged. This double-exposure photograph demonstrates clearly the correct operation of the stage, as well as emphasising the absence of a differential offset. There was no apparent change in the output signal for common-mode changes of up to 10 V at the input.

Fig. 3 demonstrates the application of the device as a practical operational amplifier and comparator. Fig. 3*b* illustrates the response of the amplifier to a triangular waveform when connected as a unity-gain voltage follower. In this example, a single-stage m.o.s. inverter having a gain of −20 was directly connected to the output of the differential stage;



682/2

Fig. 2 *Ramp response of differential stage*
Input = 2 V peak to peak, f_c = 2 MHz, output = 1 V peak to peak, $C_1 = C_2$



682/3

Fig. 3 *Response to input*
a (2 V peak to peak) f_c = 2 MHz
b Voltage-follower operational amplifier (1·9 V peak to peak)
c Comparator (5 V peak to peak)

the short-circuit reset technique was employed. Fig. 3*c* illustrates the response of the amplifier to the same waveform when connected as a comparator. In this example, a further m.o.s. inverter stage was connected at the output to boost the gain and clip the response. The clock frequency was again 2 MHz and operation of the differential stage in both operational-amplifier and comparator applications is confirmed.

Integration: The circuits presented here may be readily fabricated by using m.o.s. capacitors and transistors, so that monolithic integration of the amplifier seems quite feasible. Indeed, a direct realisation of the circuit of Fig. 1, including a 3-stage m.o.s. amplifier and an output buffer need involve only 11 m.o.s. transistors and 2 capacitors, with no requirement for matched devices. The projected layout area of such a device is $1{\cdot}6 \times 10^{-7}$ m^2, a considerable reduction over the size of a conventional circuit. Data convertors requiring banks of parallel comparators appear to be a natural application, and the device is also well suited to linearisation of the c.c.d. transfer function.[4]

Conclusions: A simple, new differential stage has been presented requiring few components and having good inherent immunity to common-mode signals and bias variations. The circuit is readily realisable in any m.o.s. technology and requires a biphase clock to provide a chopper-stabilised true differential output. The differential stage may be readily followed by conventional m.o.s. amplifiers to realise potential open-loop gains of about 60 dB.

A bias technique has been proposed for use with the amplifier which does not require the generation of external bias voltages, but which gives good immunity against process and drift variations. The projected signal bandwidth of an integrated amplifier requiring only $1{\cdot}6 \times 10^{-7}$ m^2 is 1 MHz, and operation of the circuit has been demonstrated at these frequencies by using discrete devices.

Acknowledgment: The authors gratefully acknowledge the support of AUWE, Portland and the Wolfson Microelectronics Liaison Unit, University of Edinburgh.

P. B. DENYER

Wolfson Microelectronics Liaison Unit
University of Edinburgh
Mayfield Road
Edinburgh EH9 3JL, Scotland

J. MAVOR

Department of Electrical Engineering
University of Edinburgh
Mayfield Road
Edinburgh EH9 3JL, Scotland

8th November 1977

References

1 FRY, P. W.: 'A MOST differential amplifier', *IEEE J. Solid-State Circuits*, 1969, **SC-4**, pp. 166–168
2 WESTE, N., and MAVOR, J.: 'An MOS amplifier for CCD applications'. Proceedings of 1976 international conference on the technology and applications of CCD's, Edinburgh, Sept. 1976, pp. 326–338
3 TSIVIDIS, Y. P.: 'An integrated NMOS operational amplifier with internal compensation', *IEEE J. Solid-State Circuits*, 1976, **SC-11**, pp. 748–753
4 McLENNAN, D. J., and MAVOR, J.: 'Novel technique for the linearisation of charge-coupled devices', *Electron. Lett.*, 1975, **11**, pp. 222–223

Miniature Programmable Transversal Filter Using CCD/MOS Technology

PETER B. DENYER, JOHN MAVOR, MEMBER, IEEE, AND JOHN W. ARTHUR

Abstract—This paper describes the operational features and performance of an integrated-circuit programmable sampled-analog data filter in transversal form using CCD/MOST technology. Reasons behind the particular choice of filter architecture for a prototype realization and its comparison with other reported designs in this technology are discussed, with particular emphasis placed on a novel MOST multiplier array implementation. The performance characteristics of a prototype 64-point filter design based on this approach is detailed in the context of frequency- and matched-filtering, and a module of 256 points using four cascaded filters is also described. Techniques for optimizing the inherent performance limits of these filter types under microprocessor control are suggested, via the iterative adaption of the filter impulse response, and results are given to show the improvement obtained. Finally, the potential of this miniature integrated-circuit filter for sonar type applications is briefly discussed.

I. Introduction

THE transversal filter, shown schematically in Fig. 1, is a general purpose sampled-data signal-processing element useful for matched filtering operations and for realizing all-zero responses in the frequency domain. In operation, the filter is conceptually simple—input samples are successively delayed and multiplied by a set of weighting coefficients with all products summed within each time period to form the output samples. Mathematically, we may write

$$c(n) = \sum_{m=1}^{N} s(n-m) r_m. \tag{I.1}$$

Manuscript received May 10, 1978; revised August 10, 1978 and September 12, 1978. This work was supported by the AUWE, Portland, U.K.

P. B. Denyer and J. W. Arthur are with the Wolfson Microelectronics Institute, University of Edinburgh, Edinburgh EH9 3JL, Scotland, U.K.

J. Mavor is with the Department of Electrical Engineering, University of Edinburgh, Edinburgh EH9 3JL, Scotland, U.K.

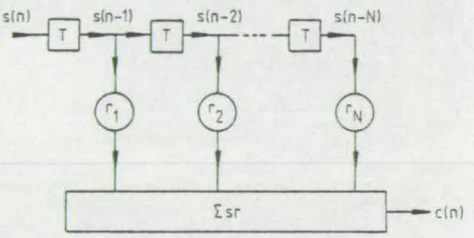


Fig. 1. Transversal filter.

This is often referred to as the convolution sum because it represents a convolution of the input sequence with the weighting coefficients which form the impulse response of the device.

The suitability of charge-transfer devices for general purpose signal processing, using the transversal filter concept, is now well recognized. Such filter realizations have thusfar focused practically on two techniques, namely the resistively weighted tapped delay line [1] and the split-gate tapped delay line [2]. A common feature of both these systems is the inflexible nature of the weighting coefficient program, and, therefore, of the associated impulse and frequency responses. Thus both systems are applicable only where a *fixed* filter response is required.

The availability of a monolithic filter with electrically *programmable* coefficients has two main attractions. First, the possibility of a compact, versatile analog filter under remote electrical control opens up many novel and independent application areas. Second, when combined with a permanent reference memory which is user-programmable, this filter forms an economically versatile alternative to split-gate devices for dedicated stand-alone applications.

0018-9219/79/0100-0042$00.75

Many different programmable filter architectures have been proposed although sufficient results with which they may be compared are not yet available. Generally two filter types may be identified—those which employ analog weighting coefficients and multipliers, and those which combine digital weighting coefficients with some form of multiplying digital-to-analog converter (DAC) at every filter point. Even within these classes different realizations are possible depending upon the form of signal and reference storage and the manner in which the necessary time-shift operation is achieved. Until recently, however, the complexity of the integrated filter structure has limited practical realizations.

Two viable approaches have been demonstrated in monolithic form—one an analog coefficient filter [3] and the other a digital coefficient filter [4]. Other devices which incorporate a single-bit digital reference memory and multipliers have been reported [5], [6]. These devices may be paralleled and their outputs suitably weighted to achieve equivalent multibit references [7].

In this paper, we present results on the operation of an analog coefficient filter [8], [9], which is thought to have several advantages over previously reported devices for both matched- and frequency-filtering applications. A full treatment of an equivalent filter with an integrated digital coefficient memory may be found in [4]. In Section II, the analog coefficient device architecture is reported, and rationalized in comparison with other realizations. Previous experience in the development of these filters [10] has shown that the multiplier elements are a critical factor; Section III deals specifically with a novel analog multiplier arrangement used in this filter realization.

II. Device Architecture

The chosen transversal filter architecture shown in Fig. 2 is both simple and compact, involving the minimum of signal manipulation and integrated silicon area. It is a *direct* realization of the block diagram of Fig. 1, and has been implemented using linear CTD and MOS component technology.

The signal register is formed from a tapped analog CCD-delay line [11] which simultaneously realizes the signal storage and time-shift operations required in the convolution sum (I.1). CTD realization of this element is optimum for these applications in that the analog time-delay and shift processes are achieved naturally, with the most economical use of silicon area. We have employed a single three-phase CCD register with delay outputs (taps) implemented using the floating-gate reset (FGR) sensing technique [9].

Because the CCD signal register provides the necessary time-shift process, a stationary[1] analog reference register is sufficient, supplying the weighting values to the multipliers in parallel form. An electrically simple and physically compact realization of this element uses discrete MOS capacitors for analog voltage memory. These feed the multiplier reference terminals via buffer amplifiers, and reference values are updated individually via a single digitally multiplexed analog input bus. Note that feedback-linearization of the reference channel, along a similarly multiplexed feedback bus, is also feasible.

This static analog reference memory implies a simpler chip structure than filters which employ both dynamic signal and reference registers. Note that it is possible to exchange the signal and reference register functions (such that the signal is loaded into the static register and the reference coefficients are time shifted and recirculated) without changing its function.

[1] Spatially "stationary" relative to the signal information.

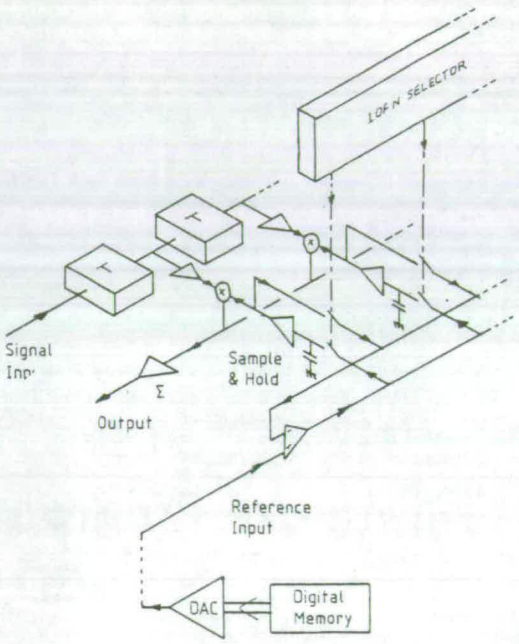


Fig. 2. Block diagram of device architecture showing external reference memory facility.

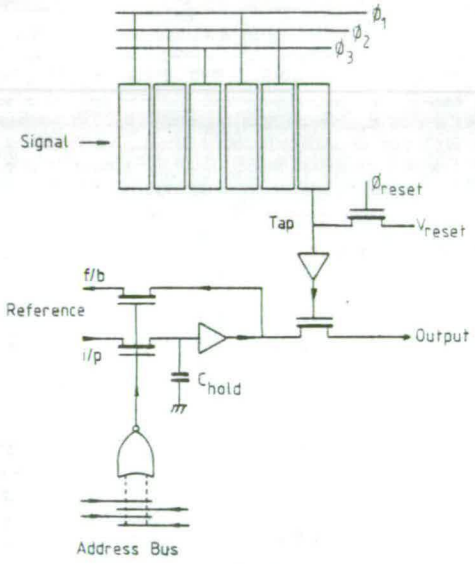


Fig. 3. Circuit diagram of one filter cell.

Unfortunately, analog coefficient storage is by nature dynamic and thus may require some form of dedicated, external memory for refresh purposes, as shown in Fig. 2. Despite this two-level memory requirement, the arrangement is still optimally compact for filters of more than (approximately) 32 points because of the relative simplicity of the analog multipliers compared with the multiplying DAC structures implied with a single digital-reference memory architecture.

For general purpose signal processing, accurate multiplication of the signal and reference (weighting coefficient) samples is required at each filter point. Now it is well known that

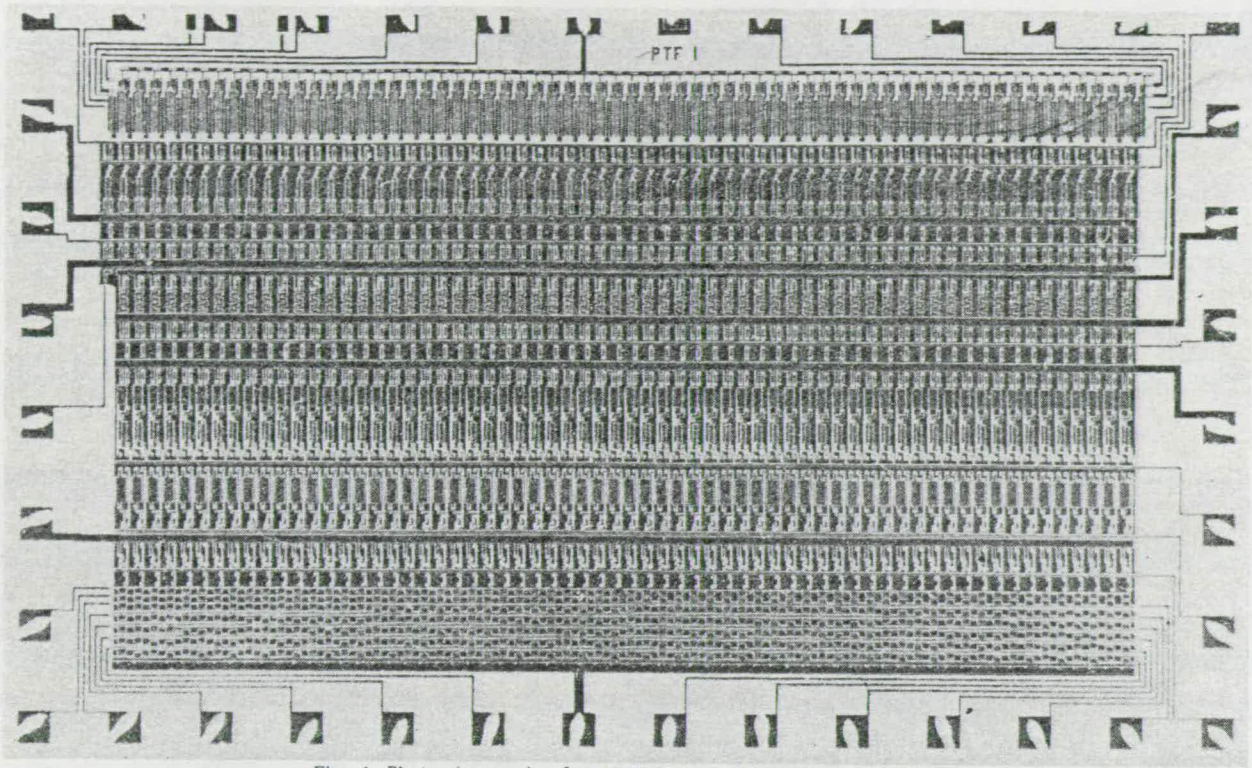


Fig. 4. Photomicrograph of prototype monolithic filter containing 64 points. Size is 184 × 136 mils.

MOS transistors, operating in the "triode" (presaturation of drain current) region, potentially offer a compact multiplying element. Realizations have suffered, however, from poor accuracy and dynamic range in addition to long-term stability and drift problems. For this reason, a multiplication arrangement has been developed around a *single* MOS transistor which gives much improved performance over contemporary realizations. In Section III, the operation and performance of this element are described in more detail.

In Fig. 3, the circuit schematic of one cell of this filter architecture is shown. The simplicity of the structure and the minimal number of components involved are of particular note. Using this arrangement we have realized a 64-point programmable transversal filter [7] with 56-μm cell pitch, fabricated in a metal-gate process [12], having a chip size of 184 mils × 136 mils. Power dissipation in the sense and buffer amplifiers, and in the address logic, is set at 300 mW (that is a value of approximately 5 mW per filter point). A photograph of the prototype chip is shown in Fig. 4.

III. Multiplication

This multiplication technique exploits the essentially linear transconductance property of the MOS transistor operating in the presaturation, or "triode" region. In this mode any *change* in gate-source voltage, ΔV_{GS}, stimulates a change in drain current which is proportional to both ΔV_{GS} and to the existing drain-source voltage V_{DS}. Formally, for the MOS transistor operating in the presaturation region, we may write (to a first approximation)

$$I_{DS} = \beta_M \left[(V_{GS} - V_T) V_{DS} - \frac{V_{DS}^2}{2} \right] \qquad \text{(III.1)}$$

where β_M is a process dependent gain constant and the other symbols have their usual meaning.

For a constant drain–source voltage V_{DS}, any change in gate voltage ΔV_{GS} stimulates a change in drain current given by

$$\Delta I_{DS} = \beta_M V_{DS} \Delta V_{GS} \qquad \text{(III.2)}$$

for all sign combinations of V_{DS} and ΔV_{GS}.[2]

This change in drain current appears on a quiescent current given by

$$I_0 = \beta_M \left([V_{GS_0} - V_T] V_{DS} - \frac{V_{DS}^2}{2} \right) \qquad \text{(III.3)}$$

where V_{GS_0} is the gate voltage corresponding to "zero" signal. All terms on the right-hand side of (III.3) are constant for a given reference coefficient, thus I_0 remains constant.

In relation to the filter circuit, we may multiply the signal and reference samples at each filter point using the MOS transistor as follows. The reference sample which remains constant is applied to the transistor drain, while the source is held at a quiescent level equivalent to the reference-zero r_0. The signal sample is applied to the transistor gate, and this change in gate voltage stimulates the required change in drain current from (III.2). Currents from all the multiplier transistors may be summed on a common source busbar (shown in Fig. 5) which is held at the required reference potential by the output summing amplifier (off-chip).

The unwanted quiescent output component I_0, may be canceled by alternatively switching zero-samples V_{GS_0} onto all

[2] If the drain and source potentials remain constant so does the threshold voltage; thus (III.2) remains true for more complex drain current expressions than (III.1), involving threshold bias dependence.

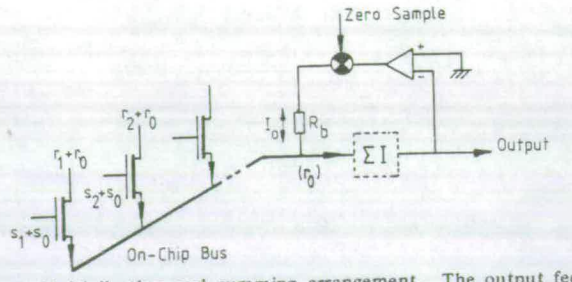


Fig. 5. Multiplication and summing arrangement. The output feedback loop removes the unwanted current component I_0.

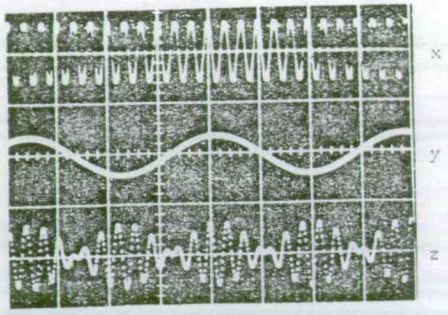
of the multiplier gates. I_0 can then be detected and cancelled at the output. As true signal samples are subsequently switched onto the multiplier gates via the CCD taps, the resultant output corresponds to the change in drain current, which is the sum-of-products term required; the quiescent current level remains cancelled. A summing circuit which achieves this objective is shown also in Fig. 5. A feedback loop is enabled during the zeroing phase to bleed the quiescent current I_0 from the summing bus itself, via the current bleed resistor R_B. Signal zeroes are conveniently switched on to the multiplier transistor gates through the CCD-tap reset transistors. Note that if any of the reference coefficients are changed, the new quiescent current I_0 is immediately detected and cancelled during the next switching operation, which may take place at the system sampling frequency.

An advantage of this time-multiplexed multiplication scheme is that the desired and the unwanted output components are generated by the same MOS transistors, thus obviating the mismatching errors associated with the dual-transistor balanced multiplier. Also variations in the transistor-threshold voltage affect only I_0, via (III.3), which is cancelled; thus they also contribute no error term.

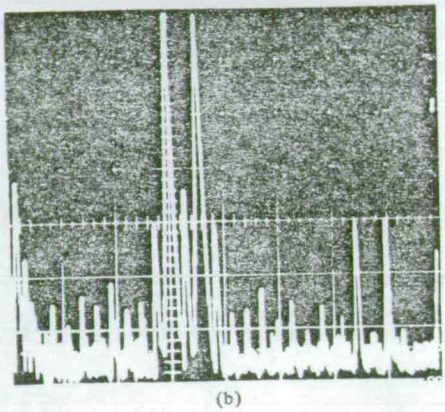
If the filter output is to be ac coupled, or the dc level is otherwise unimportant, we do not need to cancel the quiescent output component associated with I_0. The zero-switching requirement is thus removed and operation of the device is simplified. A simple summing amplifier (possibly a common-base bipolar stage) is sufficient, with no cancellation circuitry.

For general purpose testing of the prototype device we have adopted the former time-multiplexed multiplication scheme. The multiplication of two sinusoids using this technique is demonstrated in Fig. 6(a). This operation is useful in characterizing the performance of the multiplier in that all four quadrants are simultaneously exercised over the full range, allowing a meaningful analysis of harmonic content in the frequency domain. The frequency spectrum of this product is shown in Fig. 6(b), from which the expected sum- and difference-frequency components are clearly dominant. The additional spectral component patterns indicate two error terms, of the form xy^2 and x^2y, at relative amplitudes of about 2 percent with respect to the desired xy product. The significance of these error terms reduces proportionately with applied signal magnitude.

We believe that these errors are caused by a limited dependence of the gain of the multiplier transistor upon the surface field in the channel, which is directly affected by the applied gate voltage. As such the inherent multiplier distortion depends in a complex way upon the operating conditions of the device although the effect is minimized for gate drives well into the triode region. For example, doubling the quiescent



(a)



(b)

Fig. 6. Multiplication of two sinusoids using the time-multiplexed four-quadrant transconductance technique. (a) Time domain, f_x = 3 kHz (1-V p-p), $V_{GS_0} - V_{r_0} - V_T = 4$V, f_y = 300 Hz (1-V p-p), V_{r_0} = 4 V. (b) Frequency domain, 10 dB/cm: 1 kHz/cm. Note dominant sum and difference frequency components.

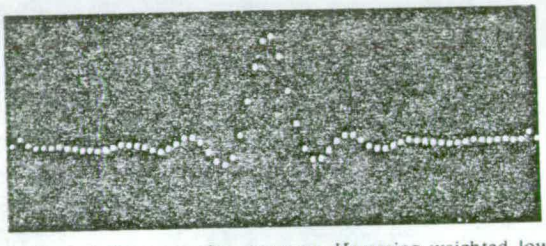


Fig. 7. Typical filter impulse response; Hamming weighted low-pass filter.

gate-drive voltage reduces the distortion components in Fig. 6 to 1 percent.

IV. Performance

We now examine the potential of this filter realization by presenting practical results obtained using the monolithic 64-point device and deriving some important performance parameters from them.

A. Dynamic Range

A stringent initial test of device accuracy and dynamic range is to examine the impulse response, a typical example of which is shown in Fig. 7. Here the reference coefficients represent a Hamming-weighted sinc x function used to realize a low-pass filter response. The impulse test is a good indication of performance because only *one* filter point is active at any time, while all filter points continue to contribute noise. The measured dynamic range of the impulse response for the 64-point filter is 34 dB. Clearly though, this parameter

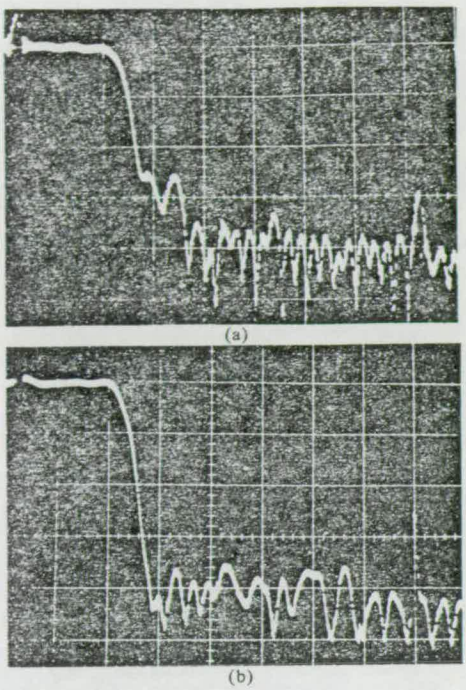


Fig. 8. Filter frequency responses with impulse response as Fig. 7. 10 dB/cm: 1 kHz/cm. (a) Directly applied reference (with linearization). (b) After-manual correction of impulse response via reference coefficients (see Section V).

is a direct function of the number of filter points used, as each additional point contributes noise. A more useful basis of comparison between different filters is to refer the dynamic range to a filter containing a single point; thus the dynamic range per filter point D_{fp} is given by

$$D_{fp} = D_{ir} + 10 \log_{10} N \qquad \text{(IV.1)}$$

where D_{ir} is the measured dynamic range of the impulse response and N is the total number of filter points.

For this device then

$$D_{fp} = 52 \text{ dB}.$$

We may use these parameters to predict the dynamic range of the device in any application by calculating the peak-attainable output signal and referring this to the known noise level. Now, the maximum possible dynamic range corresponds to the matched filtering of clipped sequences (e.g., square waves) whence the peak output dynamic range D_{op} is given by

$$D_{op} = D_{fp} + 10 \log_{10} N. \qquad \text{(IV.2)}$$

for the 64-point filter reported here;

$$D_{op} = 70 \text{ dB}$$

which has been verified by measurement.

It appears from (IV.2) that the peak filter dynamic range may be increased indefinitely by increasing N. This is not the case in practice of course, as ultimate limits are imposed by the summing amplifier. For example, the summing configuration of Fig. 5 imposes a practical achievable output dynamic range of 80 dB.

Where the reference coefficients are not fully clipped (such as is shown in Fig. 7), then the peak output signal and thus the dynamic range are reduced in proportion to the average reference power. Clearly care must be taken in implementing impulse responses having little average power.

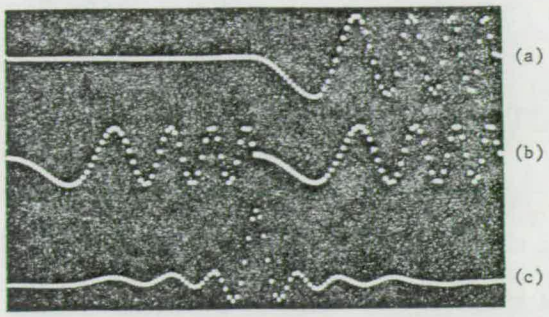


Fig. 9. Matched chirp detection $f_S = 6.4$ kHz. Chirp from 0 Hz to $f_S/8$ Hz. Trace (a) signal cascade output from 64th tap. (b) Reference waveform. (c) Filter output showing correlation peak.

B. Accuracy

The equivalent accuracy of the reference coefficients and multiplier elements may be readily estimated from the stopband performance of the low-pass filter realization.

It is possible to design ideal low-pass transversal filters having defined stopband characteristics using, for example, windowing techniques or "optimal" filter design programs [11]. Normally, the maximum stopband attenuation that can be achieved in an ideal filter increases with the number of stages used. However, if tap weight inaccuracies are present, an average normalized output error E results which also increases with N and is given by [14]

$$E^2 = \frac{(\text{TOL})^2}{3} \cdot \frac{N}{(\Sigma r)^2} \qquad \text{(IV.3)}$$

where TOL is the tolerance on the maximum tap weight.

This average, or expected output error will mask the ideal attenuation in the stopbands and, in practice, will impose a limit on realizable stopband levels. Greater attenuation can then only be achieved by cascading filter sections [14].

The effective tap weight tolerance may thus be estimated by measuring the performance of a low-pass filter realization having a large stopband attenuation. The frequency response of such a realization, using the prototype 64-point device, is shown in Fig. 8(a); the theoretical stopband level is approximately -50 dB. From this figure the average stopband attenuation achieved is about 34 dB and has a minimum value of 26 dB. From (IV.3) we calculate the effective tap weight tolerance to be 2 percent and attribute this to individual multiplier gain errors caused by oxide thickness variation and to random offset errors in the reference coefficient buffers, caused by threshold variation. The frequency response of Fig. 8(b), however, demonstrates an average stopband attenuation of 40 dB (minimum value 34 dB) with a corresponding tap weight tolerance of 1 percent. This improvement was achieved by adjusting the reference coefficients to correct for these multiplier errors and is treated in more detail in Section V.

C. Matched Filtering

A prime application of the programmable device is as a matched-filter, or correlation detector. Essentially, the impulse response of the filter is chosen to be the time-inverse of the waveform to be detected. Suitable waveforms may be chosen such that the signal energy is time compressed into a

single output peak. Fig. 9 shows the matched-filter detection of a chirp waveform, generating a correlation peak of the sinc x form. The compression factor is directly proportional to the time–bandwidth product of the chirp; and, for a filter of N points, the maximum TB product that may be achieved is

$$TB_{max} = \frac{N}{2}. \tag{IV.4}$$

These waveforms are commonly employed in sonar and radar equipments and are used to maximize transmitted signal energy while retaining range definition at the detector output. Clearly high TB figures are desirable, and may be achieved by cascading devices to increase the number of filter points. The prototype filter design is directly cascadable through the CCD-signal register, and Fig 10(a) shows a 256-point programmable filter module containing four cascaded 64-point filters. The printed-circuit board measures 6.3 × 9.2 in and dissipates 3 W, of which 1.2 W is contributed by the correlator chips.

The matched-filter detection of a chirp of TB = 128 using the 256-point module is demonstrated in Fig 10(b).

Charge transfer inefficiency (CTI) ultimately limits the number of filter points that may be cascaded with this filter configuration, as the signal degenerates along the CCD register. However, computer simulation of this effect shows that results are quite suitable for

$$TB\epsilon < 1$$

where ϵ is the CTI value per filter stage. Thus for a typical ϵ per stage of 10^{-3}, chirps of TB = 1000 may be processed, implying a total filter length of at least 2000 points from (IV.4).

D. Speed

A major bandwidth restriction in this prototype circuit occurs in the summing amplifier arrangement shown in Fig 5. The zeroing feedback-loop incorporates a commercial sample and hold unit requiring settling times up to 10 μs in practice and limiting the system clock frequency to 100 kHz. The internal chip circuitry and current summing network is potentially much faster, however, and future systems are intended to sample at 2 MHz, using an ac coupled output circuit.

V. Microprocessor Interface

The programmable aspect of these analog filters make them naturally suitable for processor control. The microprocessor, in particular, is a highly compatible computing element, allowing software programmability and intelligent choice of filter characteristics based on environmental inputs.

Apart from these general advantages, however, the microprocessor may be used with great effect to improve the accuracy of these filters via *iterative reference adaptation*. Quite simply, the actual impulse response of the filter can be compared to the desired response, and then corrected to the required degree of accuracy via control of the individual reference coefficients. Here filters employing a static reference are at a great advantage in that each multiplier is associated uniquely with one reference coefficient, and any weighting errors in the multiplier may be corrected by adjusting that coefficient.

A block diagram of the reference adaptation loop is shown in Fig. 11, the analog-to-digital converter (ADC) converts the impulse response to the required digital form; and, after taking a suitable number of averages, the microprocessor compares this with the desired value and then applies the necessary correction. The programmable filter is now updated with the ideal reference coefficients plus corrections via the DAC, and the process is iterated as required.

More formally, let us represent multiplier errors by signal and reference polynomials, so that

$$s \rightarrow (s + \alpha_0 + \alpha_1 s + \alpha_2 s^2) \tag{V.1a}$$

and

$$r \rightarrow (r + \beta_0 + \beta_1 r + \beta_2 r^2). \tag{V.1b}$$

We have already shown how multiplier errors can be identified in this way, but we may further take these polynomials to include all offset and harmonic errors within the signal and reference channels. Allowing also for gain errors G_e between multipliers, we see that

$$sr \rightarrow (1 + G_e)(s + \alpha_0 + \alpha_1 s + \cdots)(r + \beta_0 + \beta_1 r + \cdots). \tag{V.2a}$$

In fact the effects of offset errors in the signal channel α_0 are static and thus automatically cancelled using the single-transistor multiplication technique [7]; so we may write

$$sr \rightarrow (1 + G_e)(1 + \alpha_1)\left(s + \frac{\alpha_2 s^2}{1 + \alpha_1} + \cdots\right)(r + \beta_0 + \beta_1 r + \cdots). \tag{V.2b}$$

Now let the corrected reference coefficient be r and choose r (by iterative adaptation), such that

$$(1 + \alpha_1)(1 + G_e)(r + \beta_0 + \beta_1 r + \cdots) = r. \tag{V.3}$$

All of the terms in this equation are constant and thus the correction becomes permanently valid. Thus we obtain the new product

$$s_r = r\left(s + \frac{\alpha_2 s^2}{1 + \alpha_1} + \cdots\right). \tag{V.4}$$

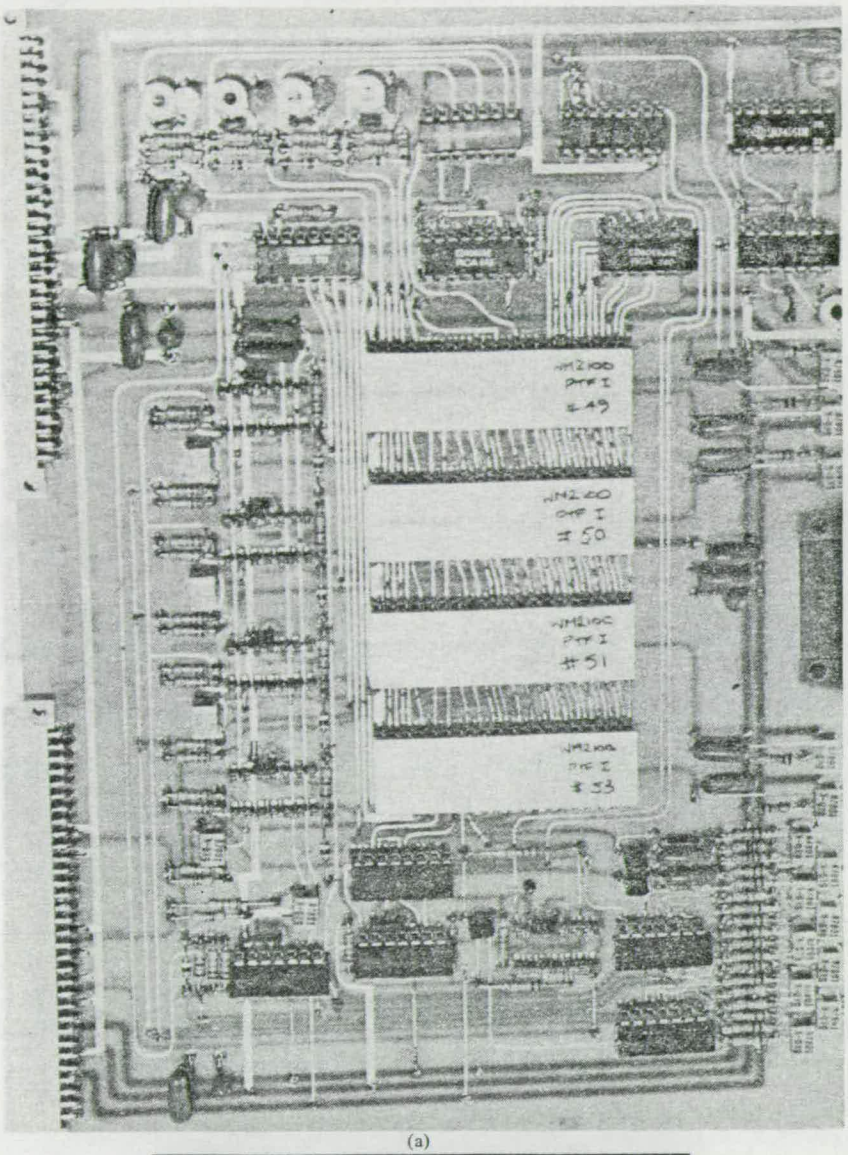
The only remaining error is that associated with harmonic distortion of the signal; the weighting factor has been made ideal[3].

Naturally the impulse response is in practice noisy, and a suitable number of responses must be averaged before applying a correction. If we assume that the noise has a normal distribution with standard deviation of σ percent, then for 95 percent confidence that the averaged response is within Ω percent of the value without noise we require

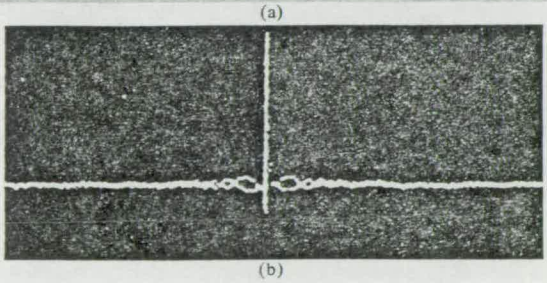
$$I = \left(\frac{2\sigma}{\Omega}\right)^2 \tag{V.5}$$

samples. For example, to improve the impulse response of the 64-point filter reported in Section IV to 8-bits accuracy (Ω = 0.2 percent) with an initial dynamic range of 34 dB (σ = 2 percent) requires an average to be taken of 400 responses, requiring approximately 250 ms at 100-kHz sample rate. The number of iterations required depends upon the form of error to be corrected, but 4 iterations appear significant in simulations involving gain errors up to 10 percent

[3]This process was implemented manually to achieve the result of Fig. 8(b) by correcting the reference coefficients until the impulse response became (visually) "ideal."



(a)



(b)

Fig. 10. 256-point programmable module. (a) Printed circuit board containing four cascaded 64-point filters and all drive electronics. (b) Matched chirp detection. Chirp from 0 Hz to $f_s/2$ Hz; TB = 128. Note sharp peak of maximum compression.

and second harmonic distortion to 3 percent. Thus using this system, we may expect to achieve tap weight accuracies of 8 bits over an adaptation period of one second.

Before leaving this topic, we identify an important alternative application area. Having adapted the filter response as required, the contents of the digital reference and correction memories may be combined and transferred to PROM. The reference channel of the filter may then be driven directly from the PROM plus DAC to form a stand-alone accurately weighted transversal filter, equivalent in many aspects to a split-gate device. Clearly the processor and adaptive loop form a single laboratory facility for servicing programmable-

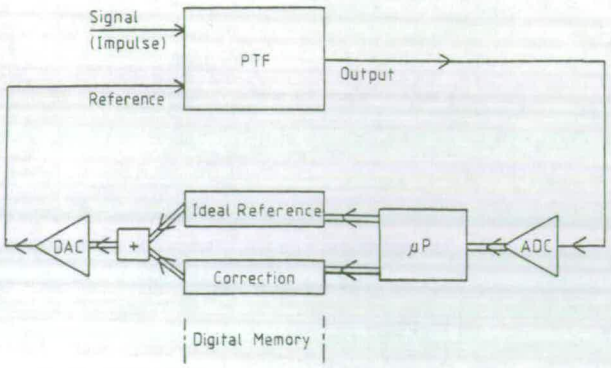


Fig. 11. Block diagram of microprocessor interface for *iterative reference adaptation.*

TABLE I
PERFORMANCE SUMMARY OF CCD/MOS MONOLITHIC TRANSVERSAL FILTER

Parameter	Prototype Result	Future Devices
Filter points per chip	64	32/64 (frequency filter) 256 (matched filter)
Total potential cascadable points (matched filtering)	1000	2000
Power consumption	300 mW/chip (5 mW/point)	1 mW/point
Dynamic range per filter point D_{fp}	52 dB	up to 60 dB
Weighting accuracy	2 percent	2 percent
with *reference adaptation*	better than 1 percent	better than 1 percent
Harmonic (signal) distortion	−34 dB	−40 dB
Storage time (reference)	4 s	up to 10 s (dependent on process)
Signal bandwidth	50 kHz	>1 MHz
Chip area	184 × 136 mils (2.2 × 120 mils/point)	1.1 × 120 mils/point

filter-plus-PROM modules, which might take circuit board, hybrid, or fully integrated form. No additional processing is required, and turn-around can be fast and economical for users requiring a dedicated filter response to a specified weighting accuracy.

VI. CONCLUSION

We have shown in this paper how innovative CCD- and MOS-design techniques may be combined to realize an integrated programmable transversal filter with application in both frequency- and matched-filtering areas. Many different device architectures are feasible and their features have been compared in general terms. In particular, we note that a combination of static and dynamic signal and reference registers are preferable in terms of circuit simplicity, and have shown how the static reference option allows weighting coefficient accuracy to be improved. Digital and analog references have been compared and the analog form has proven to be preferable for matched filters requiring many stages, due to the improvement in packing density it offers.

These filter forms offer powerful signal-processing capability, and their programmable aspect allows flexible and fast control of the filter function, either remotely or internally, via a system-based microprocessor. The number of filter stages to be included on one chip must be a decision based upon the eventual application. For frequency filtering, relatively few filter points are required and must, in any case, be minimized to avoid dynamic range restrictions. Possibly a 32- or 64-stage cascadable block is optimum. For matched-filter applications, however, many filter points may be necessary, and it becomes desirable to realize as many points on one chip as possible. It is intended that the high-packing-density structure reported here may be used to realize a cascadable 256-point chip for high time–bandwidth applications.

Clearly these filter structures are thusfar only prototypes, but the 64-point device reported here and its contemporaries are proving that good-performance parameters may be achieved in miniature form and that, with careful design, production devices must now be feasible.

In conclusion, we present in Table I performance parameters derived from characterizational measurements on the prototype device. In the second column of this table we give target specifications for future device design, involving circuit and process improvements to reduce power dissipation and chip area and to improve dynamic range.

ACKNOWLEDGMENT

The prototype filters were designed and evaluated in the Wolfson Microelectronics Institute, University of Edinburgh, and processed at the Plessey Company, Ltd., U.K.

References

[1] See for example RETICON TAD-32 Application Note, Reticon Corp., 910 Benicia Ave, Sunnyvale, CA 94086.

[2] D. D. Buss, D. R. Collins, W. H. Bailey, and C. R. Reeves, "Transversal filtering using charge-transfer devices," *IEEE J. Solid-State Circuits*, vol. 8, pp. 138–146, Apr. 1973.

[3] P. Bosshart, "An integrated analog correlator using charge-coupled devices," in *Proc. IEEE ISSCC*, pp. 198–199, 1976.

[4] Y. A. Haque and M. A. Copeland, "Design and characterization of a real-time correlator," *IEEE J. Solid-State Circuits*, vol. SC-12, pp. 642–649, Dec. 1977.

[5] J. J. Tiemann, W. E. Engeler, and R. D. Baertsch, "A surface charge correlator," *IEEE J. Solid-State Circuits*, vol. SC-9, pp. 403–409, Dec. 1974.

[6] B. E. Burke and W. T. Lindley, "New CCD programmable transversal filter," *Electron. Lett.*, vol. 13, pp. 521–523, 1977.

[7] D. D. Buss, W. H. Bailey, and A. F. Tasch, Jr., "Signal processing applications of charge-coupled devices," in *Proc. CCD 74*, Edinburgh, pp. 179–197, 1974.

[8] J. Mavor, J. W. Arthur, and P. B. Denyer, "Analog CCD correlator using monolithic MOS multipliers," *Electron. Lett.*, vol. 13, pp. 373–374, June 1977.

[9] P. B. Denyer, J. W. Arthur, and J. Mavor, "Monolithic, programmable analog CCD transversal filter," *Electron. Lett.*, vol. 13, pp. 373–374, Nov. 1977.

[10] J. G. Harp, G. F. Vanstone, D. J. MacLennan, and J. Mavor, "Analog correlators using charge coupled devices," in *Proc. CCD 75*, pp. 229–235.

[11] P. B. Denyer and J. Mavor, "Design of CCD delay lines with floating gate taps," *IEE J. Solid-State and Electron Devices*, vol. 1, pp. 121–129, 1977.

[12] K. D. Perkins and V. A. Browne, "Sub-micron gap metal gate technology for CCD's," *Microelectron. J.*, vol. 7, pp. 14–22, 1975.

[13] L. R. Rabiner and B. Gold, *Theory and Application of Digital Signal Processing*. Englewood Cliffs, NJ: Prentice-Hall, 1975.

[14] C. M. Puckette, W. J. Butler, and D. A. Smith, "Bucket-brigade transversal filters," *IEE Trans. Commun.*, vol. 22, pp. 926–934, July 1974.

UDC 621.382:621.372.54

Indexing Terms: Semiconductor devices, Large scale integrated circuits, Digital filters, Programmable transversal filters, Least-mean-square algorithm

A monolithic c.c.d. programmable transversal filter for analogue signal processing

J. MAVOR, B.Sc., Ph.D., C.Eng., M.I.E.E., M.I.E.R.E.*

P. B. DENYER, B.Sc.†,

J. W. ARTHUR, B.Sc., Ph.D., M.Inst.P.‡

and

C. F. N. COWAN, B.Sc.*

SUMMARY

This paper describes the operational features and performance of a fully-integrated programmable transversal filter (p.t.f.), using c.c.d./m.o.s.t. technology. The choice of filter architecture for a prototype realization is discussed with particular reference to a novel multiplier array implementation using a single, time-multiplexed m.o.s. transistor. The performance characteristics of a prototype, 64-point filter design based on this approach are detailed with reference to frequency- and matched-filtering. Techniques for optimizing the performance of this analogue filter structure under microprocessor control are suggested, through the iterative adaption of the filter impulse response, and equivalent results are given to show the improvement gained. An alternative technique for improving the filter characteristics which enables it to optimize the processing of signals under certain conditions has also been demonstrated. This adaptive filter configuration is based on the linear Widrow least-mean-square (W.l.m.s.) algorithm, and has been realized using the p.t.f. with minimal additional circuitry, without the requirement for a microprocessor.

A general signal-processing module of 256-points using four cascaded filters is described; and results are presented when it is used in a sonar, matched-filtering experiment. Also a 64-point adaptive filter based on a prototype p.t.f. is described and its application to inverse filtering and self-tuning filtering is demonstrated.

Finally, the potential of this miniature integrated filter for sonar-type applications is reviewed against new developments. In particular, a 256-point monolithic p.t.f. currently in development, and the concept of a dedicated adaptive filter in single chip form.

* *Department of Electrical Engineering, University of Edinburgh, Mayfield Road, Edinburgh EH9 3JL, Scotland.*
† *Wolfson Microelectronics Institute, University of Edinburgh, Mayfield Road, Edinburgh EH9 3JL, Scotland; now with Denyer-Walmsley Microelectronics, Ltd., Edinburgh, Scotland.*
‡ *Racal-MESL Ltd., Microwave Division, Newbridge, Edinburgh; formerly with WMI.*

The Radio and Electronic Engineer, Vol. 50, No. 5, pp. 213–225 May 1980

1 Introduction

The development of charge-coupled devices[1] in complex configurations has permitted many electronic systems to be produced in compact, low-power forms particularly for mobile equipments. These performance attributes can be achieved over conventional digital hardware solutions when the c.c.d. is operated as a sampled-data, analogue signal processor. However, when the c.c.d. is used as an analogue device it has to be associated with other peripheral circuits such as sample-and-holds, buffers, etc. If the complete processor is to be realized in monolithic form then it is necessary to realize all of the circuitry in a common l.s.i. technology: fortunately combined c.c.d./m.o.s.t. integrated circuit processes permit circuit design in analogue and digital forms. Many developments have been reported[2,3] in linear m.o.s. peripheral circuitry specifically for integration with c.c.d.s.

The suitability of c.c.d.s for compact signal processing based on the transversal filter concept is now established. Such filter realizations have usually been based on either resistively-weighted tapped delay lines[4] or using split-gate delay lines.[5] With both of these approaches the transversal filter impulse and frequency responses are fixed, owing to the inflexible nature of the weighting coefficient programme.

The concept of a transversal filter with electronically variable or 'programmable' impulse responses, however, has two main attractions:[6] Firstly, a flexible filter structure results which may be controlled remotely; when associated with a microprocessor the programmable transversal filter (p.t.f.) may be considered as a powerful peripheral function which can perform signal computation in hardware form at high speed due to its parallel nature. Secondly, when combined with a permanent reference memory (r.o.m.) which is user programmable this filtering sub-system forms an economical, versatile alternative to split-gate devices for dedicated, stand-alone applications, especially when a small number of different filters are required. This enables, for example, a filter manufacturer to realize a wide repertoire of filtering functions to individual customer specifications at final test, from a stock of p.t.f. devices.

Many different programmable filter architectures have been proposed although insufficient results are yet available for direct performance comparison. Generally two filter types may be identified; those which employ analogue weighting coefficients and multipliers, and those which combine digital weighting coefficients with some form of multiplying, digital/analogue converter at every filter point. Even within these classes different realizations are possible depending upon the form of signal and reference storage and the manner in which the necessary time shift operation is achieved. Practical attempts at realizing programmable filters—certainly in

0033-7722/80/050213+13 $1.50/0

digital form—have generally resulted in large, high-power, low-bandwidth processors with their performance usually restricted by the speed and power of the multipliers. Until recently, however, the complexity of integrated filter structures has limited practical realizations, essentially because of the problem of miniaturizing the multipliers. Currently, several viable approaches have been demonstrated in monolithic form; an analogue coefficient filter[7] and a digital coefficient filter[8] have been reported. Other devices have been described[9, 10] which incorporate a single-bit, digital reference memory and multipliers. These devices may be paralleled and their outputs suitably weighted, to achieve equivalent multi-bit references.[11]

In this paper we present the latest results and applications potential of an analogue coefficient 64-point p.t.f.[6, 12, 13] It has performance parameters that make it particularly suitable for sonar, instrumentation, biomedical and communications applications. When cascaded or multiplexed the performance of the basic filter may be extended, and it may be remotely controlled from a digital source. We will present here results for the p.t.f. when used for matched- and frequency-filtering, and also demonstrate a prototype adaptive filter configured with a p.t.f. and additional circuits. Consideration is given to a single chip adaptive filter which has potential application in several signal processing situations. Finally, the future of this analogue p.t.f. approach is discussed with reference to a 256-point filter currently in development.

2 P.T.F. Design

2.1 P.T.F. Architecture

The transversal filtering function may be realized in serial or parallel forms and operate on analogue or digital signals. In previous discussion[14] we have shown the advantages of a parallel configuration filter operating on all-analogue signals.

Specifically, the transversal filter, shown schematically in Fig. 1, is a general-purpose sampled-data, signal processing element useful for matched filtering operations and for realizing all-zero responses in the frequency domain. In operation, the filter is conceptually simple; input samples are successively delayed and multiplied by a set of weighting coefficients with all products summed within each time period to form the output samples. Mathematically, we may write:

$$c(n) = \sum_{m=1}^{N} s(n-m)r_m \qquad (1)$$

where s is the serial signal sequence and r is the filter impulse response sequence of N elements. This is often referred to as the convolution sum because it represents a convolution of the input sequence with the weighting coefficients which form the impulse response of the device. The chosen transversal filter architecture shown

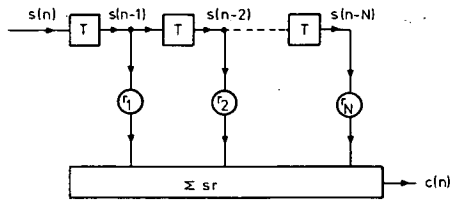


Fig. 1. A transversal filter.

in Fig. 2 is both simple and compact, involving the minimum of signal manipulation and requiring a minimum of silicon area in integrated circuit form. It is a direct realization of the block diagram of Fig. 1, and has been implemented in this work using linear c.c.d. and m.o.s. component technology.

In our design we have formed the signal register from a tapped, analogue, c.c.d. delay line[15] which simultaneously realizes the signal (eqn. (1)) storage and time-shift operations required in the convolution sum. C.c.d. realization of this element is optimum for these applications in that the analogue time-delay and shift processes are achieved inherently with the most economical use of silicon area. In a prototype design, we have employed a single, three-phase c.c.d. register with delay outputs (taps) implemented using the floating-gate, reset sensing technique.[13]

Because the c.c.d. signal register provides the necessary time-shift process, a stationary† analogue reference register is sufficient, supplying the weighting values to the multipliers in parallel form. An electrically simple and physically compact realization of this element uses discrete m.o.s. capacitors for analogue voltage memory. These feed the multiplier reference terminals, via buffer amplifiers, and reference values are updated individually via a single, digitally-multiplexed, analogue input bus.

Unfortunately, analogue coefficient storage is by nature volatile and thus may require some form of dedicated, external memory for refresh purposes, as shown in Fig. 2. Despite this digital memory requirement, the arrangement is still optimally compact for filters of more than (approximately) 32 points, because of the relative simplicity of the analogue multipliers compared with the multiplying d.a.c. structures implied with a single, digital reference memory architecture.

2.2 Multipliers

The disadvantage of parallel realizations of transversal filter over serial designs is that N multipliers are required for an N-point filter, whereas serial configurations employ only one. When the objective is a fully integrated p.t.f. then the requirement for many integrated multipliers, one at each filter point, is in itself a formidable problem.[16] The prototype p.t.f. described in

† Spatially 'stationary' relative to the signal information.

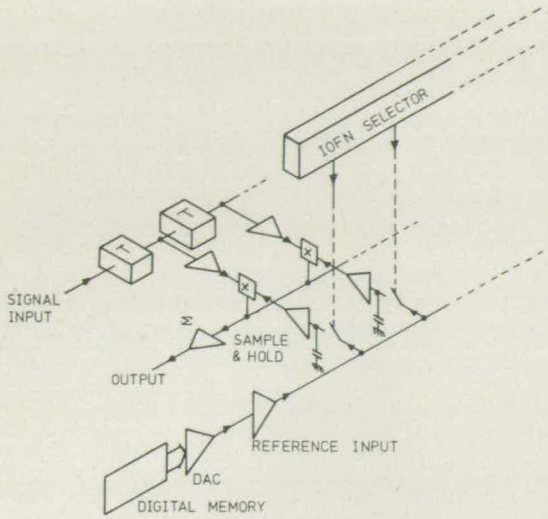


Fig. 2. The programmable transversal filter.

this paper relies on a novel analogue multiplier arrangement[6, 12] that has been developed in this work to provide accurate multiplication of the signal and reference (weighting coefficient) samples.

An economical multiplication technique based upon the essentially linear transconductance of an m.o.s. transistor operating in the 'triode' (pre-saturation of drain current) region is adopted here. A first-order expression for the drain current of such a transistor is given by:

$$I_{DS} = \beta_M \left[(V_{GS} - V_T)V_{DS} - \frac{V_{DS}^2}{2} \right] \quad (2)$$

where β_M is a process dependent gain constant and the other symbols have their usual meaning. Note that this expression does contain a potential linear multiplication term $V_{GS}V_{DS}$ but this must be isolated from the other undesirable terms.

For a constant drain-source voltage, V_{DS}, any change in gate voltage, ΔV_{GS}, stimulates a change in drain current given by:

$$\Delta I_{DS} = \beta_M V_{DS} \Delta V_{GS} \quad (3)$$

for all sign combinations of V_{DS} and V_{GS}.† Clearly this is the desired product term. This change in drain current appears on a quiescent current given by:

$$I_0 = \beta_M \left[(V_{GS_0} - V_T)V_{DS} - \frac{V_{DS}^2}{2} \right] \quad (4)$$

where V_{GS_0} is the gate voltage corresponding to 'zero' signal. All terms on the right-hand side of equation (4) are constant for a given reference coefficient, thus I_0 remains constant. The following multiplication techniques rely on cancelling I_0 to leave the pure product term given by equation (3) as output.

† If the drain and source potentials remain constant so does the threshold voltage; thus equation (3) remains true for more complex drain current expressions than equation (2) which involves threshold bias dependence.

Previous realizations of multiplier have been synthesized using two identical transistors with a common diffused terminal, which have suffered from poor accuracy and dynamic range in addition to long-term stability and drift problems. The technique has relied on the two transistor gains, β_M, being matched so that I_0 may be cancelled adequately. For this reason, in our prototype p.t.f. a multiplication arrangement has been developed around a single m.o.s. transistor which gives much improved performance over contemporary realizations. In relation to the filter circuit, we may multiply the signal and reference samples at each filter point using the m.o.s. transistor as follows. The reference sample, which remains constant, is applied to the transistor drain, whilst the source is held at a quiescent level equivalent to the reference-zero, r_0. The signal sample is applied to the transistor gate, and stimulates the required change in drain current, from equation (3). Currents from all the multiplier transistors may be summed on a common source busbar Σ, shown in Fig. 2, which is held at the required reference potential by the output summing amplifier (off-chip).

The unwanted quiescent component, I_0, may be cancelled from the output by alternatively switching zero-samples (V_{GS_0}) onto all of the multiplier gates. I_0, which contains all of the unwanted terms, can then be detected and cancelled at the output. As signal samples are subsequently switched onto the multiplier gates, via the c.c.d. taps, the resultant output corresponds to the change in drain current given by equation (3), which is the sum-of-products term required; the quiescent current level, I_0 in equation (4), remains cancelled. Provided that the reference voltages applied to the drains of the multiplying transistors remain constant during this switching operation, all of the unwanted linear and quadratic terms are exactly cancelled.

An advantage of this time-multiplexed multiplication scheme is that the desired and the unwanted output components are generated by the same m.o.s. transistors, thus obviating the mis-matching errors associated with the dual-transistor balanced multiplier. Also variations in the transistor threshold affect only I_0 via equation (4) which is cancelled, thus they also contribute no error term.

2.3 P.T.F. Implementation

A 64-point, prototype p.t.f. has been designed based on the principles discussed in Sections 2.1 and 2.2 and is illustrated in Fig. 3. Each filter point required 14 m.o.s. transistors and 2 c.c.d. bits of storage (6 gates for a 3-phase c.c.d. process). The m.o.s. peripheral circuitry required a 56 μm cell pitch when laid out in integrated circuit form and permitted two c.c.d. shift bits in this length (the tapped c.c.d. was thus 128 bits long). The device was fabricated in an n-channel metal-gate process[17] having a chip size of 4·5 mm × 3·3 mm,

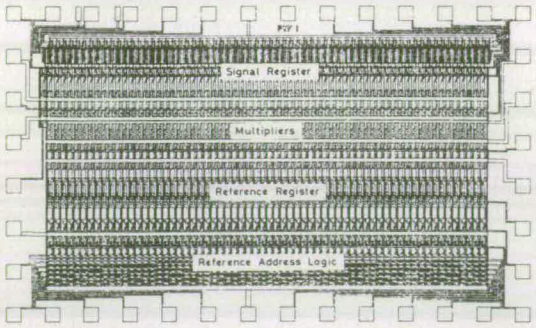


Fig. 3. The prototype 64-point p.t.f. device.

demonstrating the high packing density achievable with this design technique.

The summing amplifier was not integrated with the p.t.f. for three main reasons; although we have had experience[2] in integrating m.o.s. linear circuits to meet this requirement. Firstly, only a simple summing amplifier is sufficient because the single multiplier m.o.s.t. array requires only a single summing busbar (Σ). We have used a single, off-chip bipolar transistor stage in our experimental work which can easily be mounted near to each p.t.f. on a printed-circuit board. (Previous realizations of multiplier incorporating two matched m.o.s.t.s have required cancellation circuitry to remove undesired d.c. offsets by using three differential amplifier stages.) Secondly, we had a design priority for optimizing the number of filter points at the expense of any strictly unnecessary peripheral circuitry. Thirdly, there was a strong requirement to reduce on-chip power dissipation to minimize dark current generation, which is an exponential function of temperature, caused by chip power dissipation. Power dissipation in the sense and buffer amplifiers, and in the address logic, was set at 300 mW in the prototype 64-point p.t.f. (that is a value of approximately 5 mW per filter point).

2.4 P.T.F. Error Analysis

The prototype filter of Fig. 2 is a complex analogue structure, consequently, its error analysis is extremely involved. Here we will only identify the main imperfections in the p.t.f. and discuss their combined effect on the overall filtering function. In addition, in Section 3, where we present results for the prototype p.t.f., some discussion on the overall effect of device error will be given. However, unless precise information about both the signal and reference waveforms is available then a quantitative error analysis for a particular device is not feasible.

2.4.1 Error sources

Charge transfer inefficiency (ε) in the c.c.d. register is known to degrade signal information at high frequencies, this being a cumulative effect with increasing $N\varepsilon$ products. In terms of frequency filtering applications, however, the net effect at low frequencies is a shift in the transition edges between pass- and stop-bands of a factor $(1-\varepsilon)$, regardless of the number of filter stages.[18] The effect upon p.t.f. performance is somewhat more signal dependent but charge-transfer inefficiency is known to cause a slump in the correlation peak and a corresponding increase in sidelobe significance. It has been suggested[5] that in certain correlation applications an $N\varepsilon$ product of 2 is tolerable.

Random gain errors in the multiplication process are analogous to tap weight errors in split-gate filters and as such may be expected to impair stop-band suppression in frequency filtering applications. Correlation applications are more tolerant to these errors which become attenuated by the 'processing gain' ($N^{\frac{1}{2}}$) of the filter. The requirements placed on filter accuracy are dealt with in more detail in Section 2.4.2.

Quiescent offset errors in both the signal and reference channels appear as such at the convolver output for suitably large time-bandwidth products. Taking for example quiescent reference errors, Δr, the net output error will be $\bar{s} \cdot \overline{\Delta r}$, where the bar denotes an average value. Clearly, such errors may be reduced or eliminated where either sequence can be chosen to have zero mean value.

The finite output admittances of the buffer stages which drive the multipliers result in a signal distortion term which unfortunately precludes a general linear analysis of the resultant error. However, in correlator applications additional correlation peaks are to be expected whenever a match is detected between the generated harmonic signal component and the reference waveform.

The refreshing and subsequent decay of reference values results in a modulation of the output waveform. As this process is not generally synchronized with the incoming data sequences it appears as a form of noise at the output. Naturally, the magnitude of this noise is directly related to the decay rate and update frequency of the reference samples; these factors determine the maximum number of reference points which may be updated sequentially. Thereafter reference refreshing must take place in parallel blocks. In our experiments with the prototype filter we have found this refresh noise to be insignificant.

Although several potential error sources have been identified here it is clear that not all of them will be applicable to any one application of the device. In general, frequency filtering applications requiring large stop-band attenuations are more sensitive to these errors. Matched filtering (correlation) applications, however, are considerably more tolerant to random errors which become attenuated by the 'processing gain' of the filter; generally, the only significant errors are those which correlate with either the signal or reference sequences.

For frequency filtering applications accuracy of tap weight setting is crucial. Normally, the maximum stop-band attenuation that can be achieved in an ideal filter increases with the number of stages used. However, if tap weight inaccuracies are present, an average normalized output error, E, results which also increases with N and is given by:

$$E^2 = \frac{T^2}{3} \cdot \frac{N}{(\Sigma r)^2} \qquad (5)$$

where T is the tolerance on the maximum tap weight. This average, or expected, output error will mask the ideal attenuation in the stop-bands and in practice imposes a limit on realizable stop-band levels: greater attenuation can then only be achieved by cascading filter sections. In Section 3.1 we will discuss the effective tap weight tolerance for our p.t.f. by estimating the stop-band attenuation of a low-pass filter.

2.4.2 Error correction

The inherent inaccuracies in the individual multiplier elements can be compensated for by employing an external control loop. With this technique, which we have called iterative reference adaption, the actual response of the filter can be compared to its desired response and then corrected for using an external store. Programmable transversal filters employing a static reference are here at a great advantage in that each multiplier is associated uniquely with one reference coefficient and any weighting errors in the multiplier may be corrected by adjusting that coefficient.

The means for generating and storing the error 'signature' for an individual p.t.f. can be performed in several ways. Basically the impulse response could be adjusted manually by comparing the experimentally observed impulse response with a theoretical value and then compensating the reference input to allow for errors in a p.t.f. In Fig. 5 we have in fact applied this procedure to obtain the result (5(b)). However, in practice it would be desirable to compensate automatically for such errors in the impulse response. When applied continuously, this would have the added advantage of tracking device errors with ageing and temperature variation.

Two forms of automatic error compensation have been applied in our work: microprocessor control and

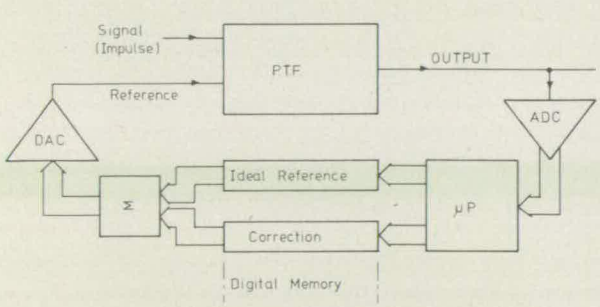


Fig. 4. A microprocessor controlled reference adaption loop.

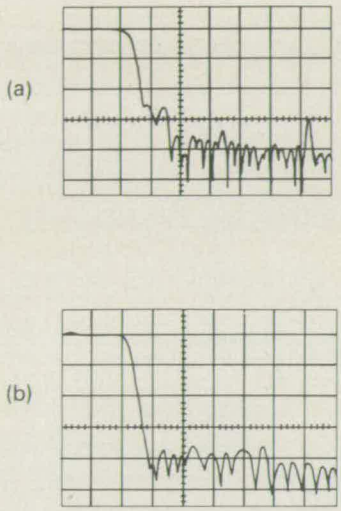


Fig. 5. (a) Frequency response of the filter when it is set up as a low-pass filter with no error compensation. (b) Frequency response when filter errors are compensated.

adaptive control using the convergence properties of the least-mean-square (l.m.s.) algorithm.[19] Both techniques employ feedback compensation; however, the adaptive filter approach does not in fact require a microprocessor. When a microprocessor is available with the p.t.f., then its software programmability makes the p.t.f.-based system very powerful and error compensation can be regarded as an inherent bonus.

A block diagram of the reference adaptation loop using microprocessor control is shown in Fig. 4. The analogue-digital converter (a.d.c.) converts the impulse response to the required digital form and, after taking a suitable number of averages, the microprocessor compares this with the desired value and then applies the necessary correction. The programmable filter is now updated with the ideal reference coefficients plus corrections, via the d.a.c., and the process is iterated as required. We have calculated[6] that to improve the impulse response of our 64-point p.t.f. to the equivalent of 8-bits digital accuracy with an initial dynamic range of 34 dB requires an average to be taken of 400 responses. This requires approximately 250 ms at a 100 kHz sample rate. The number of iterations required depends upon the form of error to be corrected but four interations appear significant in simulations involving gain errors up to 10% and second harmonic distortion to 3%. Thus, using this system, we may expect to achieve tap weight accuracies of 8 bits over an adaptation period of one second.

The alternative technique for improving the p.t.f. accuracy using the l.m.s. algorithm is discussed with reference to results in Section 3.3.

2.5 P.T.F. Performance Summary

Characterization of a general signal processing sub-system such as a p.t.f. integrated circuit is extremely complex and is dependent upon an appropriate choice of

Table 1
Performance summary of c.c.d./m.o.s. monolithic transversal filter

Parameter	Prototype result	Future device
Filter points per chip	64	32/64 (frequency filter) 256 (matched filter)
Total potential cascadable points (matched filtering)	1000	2000
Power consumption	300 mW/chip (5 mW/point)	1 mW/point
Dynamic range per filter point D_{fp}	52 dB	up to 60 dB
Weighting accuracy with reference adaptation	2% better than 1%	2% better than 1%
Harmonic (signal) distortion	−34 dB	−40 dB
Storage time (reference)	4 s	up to 10 s (dependent on process)
Signal bandwidth	50 kHz	>1 MHz
Chip area	4·5 × 3·3 mm (0·055 × 3 mm/point)	(0·0275 × 3 mm/point)

signal and reference test signals. In Table 1 we present a performance summary of results for the prototype 64-point p.t.f., and results for future device design.

The present 64-point device can process 50 kHz signals within a total chip power dissipation of 300 mW. The reference signal can be loaded into the p.t.f. at this rate and stored dynamically for up to about 4 s (thus complete reference refresh must occur on this timescale). The equivalent weighting accuracy of the basic p.t.f. is 2%, but with reference adaption as described in Section 2.4.2. this figure may be improved to better than 1%.

The measured dynamic range of the impulse response for the 64-point filter is 34 dB. Clearly though, this parameter is a direct function of the number of filter points used, as each additional point contributes noise (originating in the quiescent current of each multiplier transistor). A means of making a direct comparison between filters of any number of points is to refer the dynamic range to a filter containing a single point; thus the dynamic range per filter point, D_{fp}, is given by:

$$D_{fp} = D_{ir} + 10 \log_{10} N \qquad (6)$$

where D_{ir} is the measured dynamic range of the impulse response and N is the total number of filter points. For this device then, $D_{fp} = 52$ dB. Over this dynamic range the harmonic distortion in the signal register is better than −34 dB.

Further, for a given signal processing application, the dynamic range of the filter is a function of the processing gain for that application. Referred to peak output amplitudes with an impulse response at 0 dB, the processing gain for unweighted sine-wave autocorrelation or linear f.m. matched filtering is $20 \log_{10} N - 6{\cdot}02$ dB, and the maximum gain is $20 \log_{10} N$ dB for the autocorrelation of squarewaves. For linear f.m. matched filtering applications, the dynamic range of this device may therefore be quoted as 64 dB.

3 P.T.F. Applications

The applications potential of the prototype p.t.f. described in Section 2 is very extensive and already we have demonstrated a number of key filtering functions in compact, low-power form. The flexible nature of the p.t.f. makes it invaluable in many signal processing applications, especially for mobile equipment. We will discuss here three significant applications in which we have put it to use; although of course many others exist and are currently under investigation.

Before we turn our attention to the programmable nature of the p.t.f. it is worth discussing its use as an individual, accurately weighted transversal filter, equivalent in many aspects to a split-gate device.[5] This can be achieved by associating one or more p.t.f. chips with the 'overhead' of a dedicated digital reference store (p.r.o.m.) and d.a.c. circuit. Having adapted the filter response as required, for example using the microprocessor compensation scheme of Section 2.4.2., the contents of the digital reference and correction memories may be combined and transferred to p.r.o.m. The reference channel of the filter may then be driven directly from the p.r.o.m. plus d.a.c. to form a filter module with one or more dedicated impulse responses dependent upon the size of the memory. Clearly the processor and adaptive loop form a single laboratory facility for servicing programmable-filter-plus-p.r.o.m. modules, which might take circuit-board, hybrid, or fully integrated form. No additional processing is required and turn-around can be fast and economical for users requiring a dedicated filter response to a specified weighting accuracy. The overhead of the p.r.o.m. may be insignificant in complex signal processing situations, e.g.

a chirp-Z transform realization, where considerable memory would normally be associated with the transversal filters to perform the overall filtering function.

3.1 Frequency Filtering

By adjusting the impulse response of a p.t.f. appropriately, a frequency filter of desired response can be achieved within certain accuracy limits. For example, it is possible to design an ideal low-pass transversal filter having defined stop-band characteristics using windowing techniques or 'optimal' filter design programs.[20] However, we have seen in Section 2.4.2 that tap weight inaccuracies in practical filters cause an average output error that will mask the ideal attenuation in the stop-bands and in practice impose a limit on realizable stop-band levels. Greater attenuation can then only be achieved by cascading filter sections.[21] The effective tap weight tolerance may thus be estimated by measuring the performance of a realized low-pass filter having a large ideal stop-band attenuation.

The frequency response of such a realization, using the prototype 64-point device, is shown in Fig. 5(a); the theoretical stop-band level in this example is approximately -50 dB. From this figure the average stop-band attenuation achieved is about 34 dB and has a minimum value of 26 dB. From equation (5) we calculate the effective tap weight tolerance to be 2% and attribute this to individual multiplier gain errors caused by oxide thickness variation, and to random offset errors in the reference coefficient buffers caused by threshold variation. By visually adjusting the reference coefficients—displayed on an oscilloscope—to correct for these multiplier errors, as suggested in Section 2.4.2, an improvement in frequency response, given in Fig. 5(b), was achieved, which demonstrates an average stop-band attenuation of 40 dB (minimum value 34 dB), with a corresponding tap weight tolerance of 1%.

The advantage over split-gate filters is, of course, that filters of an experimental nature, or for applications for which no hard and fast design rules exist, can be individually specified and produced. Once a filter design is established, there is no severe obstacle to reasonable production volumes of individually optimized filters.

3.2 Matched Filtering

A prime application of the programmable device is as a matched filter, or correlation detector. Essentially, the impulse response of the filter is chosen to be the time-inverse of the waveform to be detected. Suitable waveforms may be chosen such that the signal energy is time compressed into a single output peak. Waveforms commonly employed in sonar and radar equipments are called 'chirps' (linearly modulated f.m. waveforms) and are used to maximize transmitted signal energy whilst retaining range definition at the detector output. Chirp waveforms generate a correlation peak of the sinc x form having a compression factor which is directly proportional to the time-bandwidth product of the chirp and, for a filter of N points, the maximum TB product that may be achieved is:

$$TB_{\max} = \frac{N}{2} \tag{7}$$

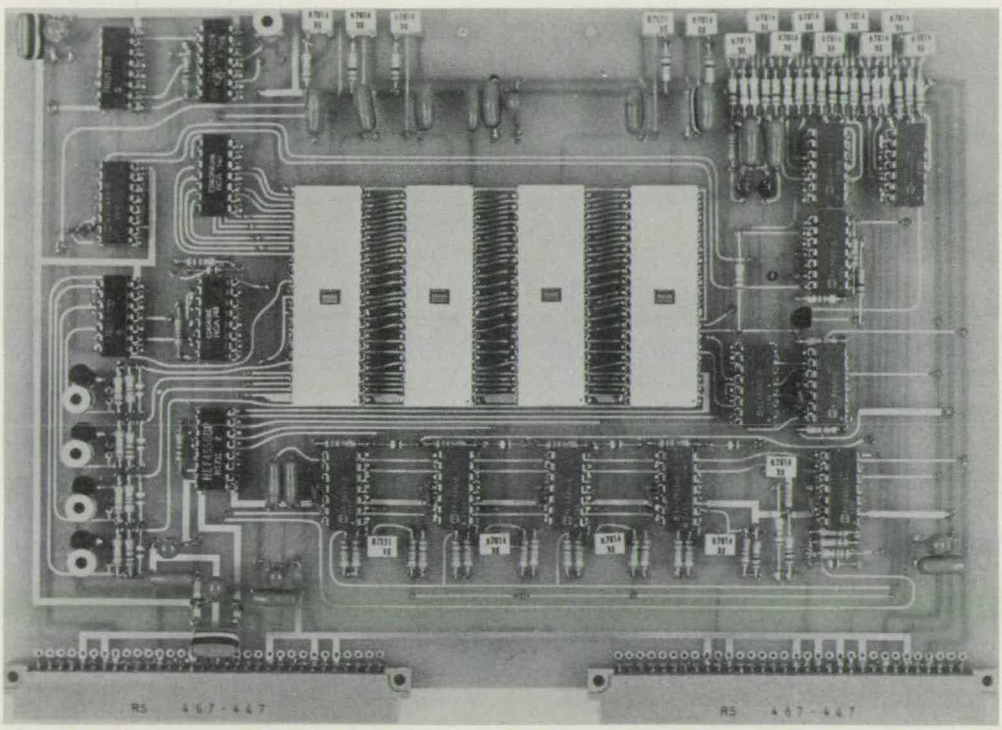


Fig. 6. A 256-point correlator system constructed by using four cascaded 64-point p.t.f. devices.

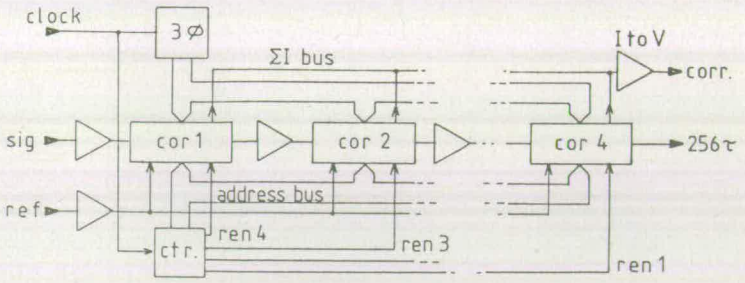


Fig. 7. The 256-point correlator system.

Clearly high *TB* figures are desirable and may be achieved by cascading p.t.f.s to increase the number of filter points.

To demonstrate the potential for the application of our c.c.d. p.t.f. to high time-bandwidth product, low-power, sonar matched-filter signal processing systems, a printed-circuit board containing four cascaded correlator i.c.s and virtually all the ancillary circuitry required to generate a complete 256-point correlator sub-system, has been produced and is shown in Fig. 6. The system configuration is given in Fig. 7. A single interstage sample-hold amplifier is required to cascade the c.c.d. delay line stages, and the reference is cascaded just as easily through a common address bus and individual r.en. (write enable) strobes. All clocking and timing circuity is provided so that the correlator samples in synchronism with an externally applied TTL clock. Externally, the correlator sub-system appears as a simple three-port element, with signal reference as inputs and correlation as output.

The complete p.t.f. sub-system is mounted on a double-sized Eurocard printed-circuit board and dissipates about 3 W, of which 1·2 W is contributed by the p.t.f. chips. Signals from d.c. to above the audio range are correlated very adequately over the full bandwidth available at a given sampling rate, giving a very versatile system. In particular f.m. matched filtering suitable to sonar signal processing may be readily accomplished. This is demonstrated in Fig. 8 which shows the performance of such a matched filter utilizing a linear f.m. sweep of maximum bandwidth, i.e. d.c. to the Nyquist limit. Performance is very acceptable considering the greater size, power consumption and cost of comparable digital processors.

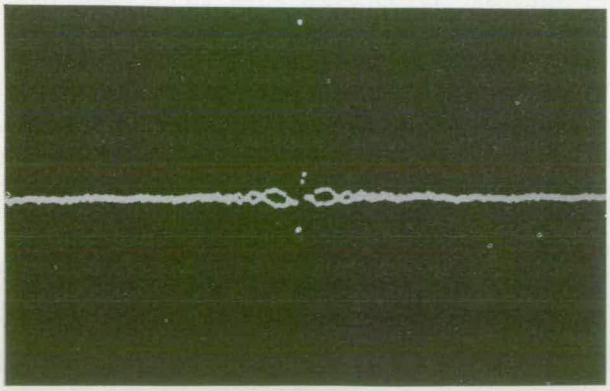


Fig. 8. Output of the 256-point correlator system configured as a matched filter to a linear frequency modulated signal swept from d.c. to the Nyquist frequency.

3.2.1 Sonar application

Two of the above 256-point correlator sub-systems have been configured into a low-i.f., 2-channel quadrature sonar processing system, Fig. 9(a), and this has been demonstrated against a submarine target in very shallow water (approximately 10 m depth), as illustrated in Fig. 9(b). The processor output shows a high resolution return in which the target is clearly visible against a background of strong reverberation.

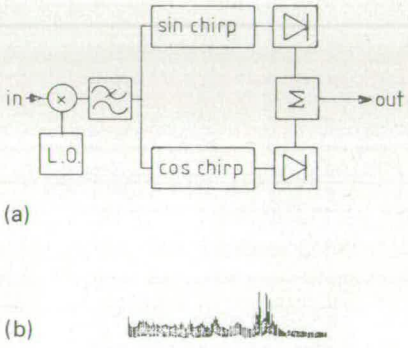


Fig. 9. (a) Low-i.f., two-channel quadrature sonar processing system using two 256-point correlator systems. (b) High-resolution sonar return from a submarine target.

3.3 Adaptive Filtering

A useful device for many applications in signal processing would be a filter needing the minimum *a priori* information about an incoming signal immersed in noise to detect and reproduce it. The adaptive filter is probably the optimum form for this type of application and it has been extensively studied in theory and modelled on computers. However, little is known about the actual physical implementation of this class of filter. The purpose of our work here is to demonstrate the feasibility of employing a p.t.f. as a central element with ancillary circuitry to illustrate basic adaptive filter operation. The eventual goal of this approach is a fully-

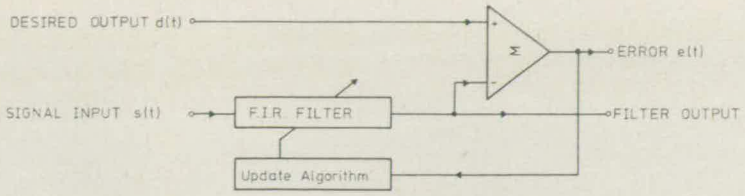


Fig. 10. Basic block diagram of an adaptive filter.

integrated adaptive filtering system for use in a wide range of applications.

3.3.1. Implementation of l.m.s. algorithm

A block diagram of the basic adaptive filter element is shown in Fig. 10. The system is supplied with two inputs; the input signal $s(t)$ and the desired filter response to this input, $d(t)$. The aim is to force the output of the filter to resemble, as closely as possible, the desired response $d(t)$. To achieve this the filter weight vector **H** is updated continuously using the following algorithm:[19]

$$h_k(t+1) = h_k(t) + 2\mu\varepsilon(t)s(t-k) \qquad (8)$$

where $\varepsilon(t) = d(t) - r(t)$ is the output error, $r(t)$ is the actual filter output and μ is a selected convergence factor. The subscript k refers to the weight or tap position in the filter, the terms in brackets are the time or delay indices, and t is in normalized units of one delay time. This is known as the Widrow least-mean-square adaption algorithm, and a detailed derivation and discussion may be found in Reference 19.

Figure 11 shows a block diagram of the system used to implement this algorithm. The circuit function within the dotted box was achieved simply by utilizing a prototype 64-point p.t.f. described in Section 2.

Since with this filter it is impossible to update all 64 weights in the time between successive output data points, the output error $\varepsilon(t)$ is sampled at only one point in 65 and the weights are updated sequentially. In addition, it is not possible to read individual weight values once they have been stored in the filter holding register. It was therefore necessary to provide a separate memory (external to the p.t.f. reference analogue store), which could be both read from, and written to, at any time to provide storage of weight values during the weight-vector updating cycle.

In the circuit implementation shown in Fig. 11 the weight store and updating loop was configured using digital circuitry with appropriate analogue-digital conversion (a.d.c.) after the convergence multiplier, and a digital-analogue converter (d.a.c.) at the input to the p.t.f. analogue store. An alternative implementation involved off-chip analogue storage based on 64 sample-and-hold stages as a precursor to a fully-integrated version based on all-analogue circuitry. This latter arrangement gave the adaptive filter an inferior performance over the digital update version essentially because it suffered from weight value decay. For certain filtering functions, in particular inverse filtering, permanent storage of the reference is required and thus digital memory would be mandatory. However, in monolithic adaptive filter designs, now at the planning stage, the all-analogue approach will be pursued further because of the improved characteristics of on-chip analogue storage (by a factor of 100 to $\sim$1 V/s).

3.3.2 Adaptive filter results

We report here results for the prototype adaptive filter based on Fig. 11 using (a) an external digital update memory[22] (as shown), and (b) an analogue memory[23] (by omitting the a.d.c. and d.a.c. and replacing the r.a.m. with an analogue unit). In Fig. 12 we show a photograph of the completed filter with an analogue update store which has been assembled on one double Eurocard-sized printed-circuit board and consumes less than 10 W of power. An additional board containing the a.d.c., d.a.c.

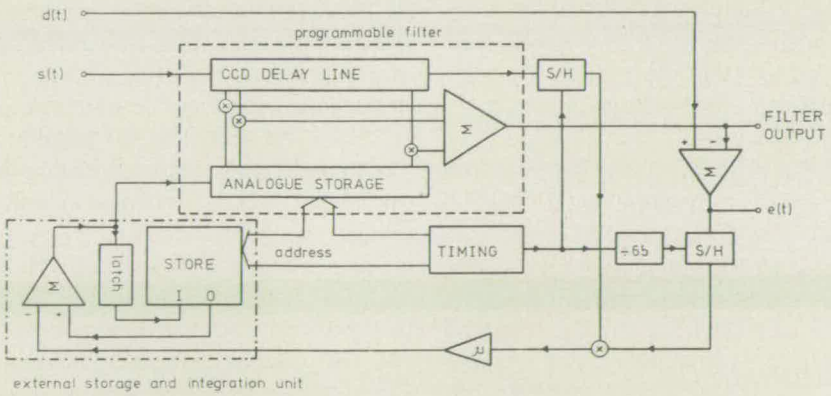


Fig. 11. Adaptive filter system using one 64-point p.t.f.

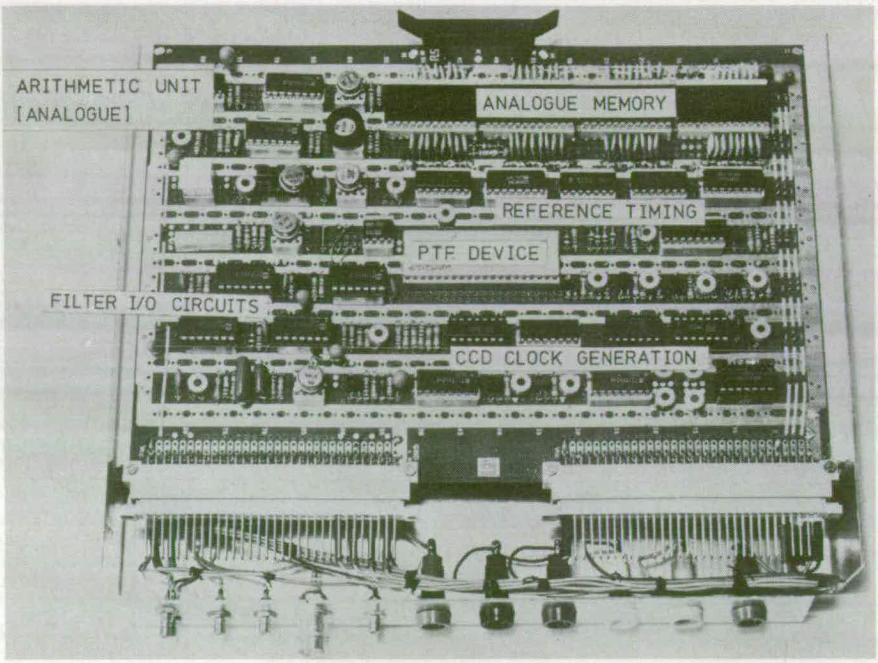


Fig. 12. The analogue adaptive filter module.

and memory is required for the analogue/digital hybrid version.

The adaptive filter unit may be used in a number of ways dependent upon the configuration of the two input ports and which of the two output ports is used. Figure 13 shows the three main filter configurations which have been used with our prototype system. Figure 13(b) shows the necessary input conditions for a noise canceller where the A input is a signal with additive noise and the B input is a correlated (but not identical) version of the noise. The filter converges to reconstitute the interfering noise at the filter output therefore subtracting coherently from A and leaving the uncorrupted signal at the error or canceller output C. One example of this type of operation would be where input A is supplied from a microphone which picks up a speech signal and some arbitrary background noise. Input B would be supplied from a second microphone which picks up only the background noise in the room. Output C should then be the uncontaminated speech signal.

As an inverse filter it is desired that the system should reconstitute a signal which has been subjected to distortion by some intervening medium, such as a transmission line. Figure 13(c) shows the adaptive filter configured as such a system. The B input is the distorted signal input and the A input is a training signal which is the required filter output. The output used in this case is D which should converge to the same waveform as A, after an initial training period during which the input signal B is known. In this time the filter impulse response converges to the inverse of the transmission line impulse response. The adaption algorithm is then switched off, freezing the weight vector, and the system is thereafter used as a straight forward transversal filter equalizer. Such a system has been described by Corl,[25] although his implementation uses a computer to update the weight vector using a zero-forcing algorithm.

The last configuration to be considered here is that of a self-tuning filter shown in Fig. 13(d); this structure is similar in operation to the noise canceller. The signal input will commonly comprise two components, (i) a narrow-band periodic component and (ii) a broadband non-periodic component. The delay T causes the non-periodic component to be decorrelated between A and B so that the filter converges to form a band-pass impulse response which produces the narrow-band periodic component of A at the filter output D (rejecting the broadband non-periodic component). Since this output is coherent in phase and amplitude with the periodic component of A, the C output then consists only of the broadband non-periodic component (periodic noise cancellation). The system effectively separates periodic and non-periodic signal components and may be used, for example, to cancel unwanted hum in speech signals and also as a self-tuning filter to reduce unwanted broadband noise on a required periodic signal. This structure has been described in detail by Widrow *et al.*[19]

As a significant and illustrative demonstration of the operation of our prototype adaptive filter with digital update, we here describe its performance when it is desired to cancel a very strong first harmonic interference on a wanted sinusoidal signal. This represents operation in the noise cancellation mode (c.f. Fig. 13(b)). Figure 14(a) shows the $s(t)$ input, i.e. the correlated interference; Fig. 14(b) is the conditioning signal input $d(t)$, in this case the input signal with interference; Fig. 14(c) is the

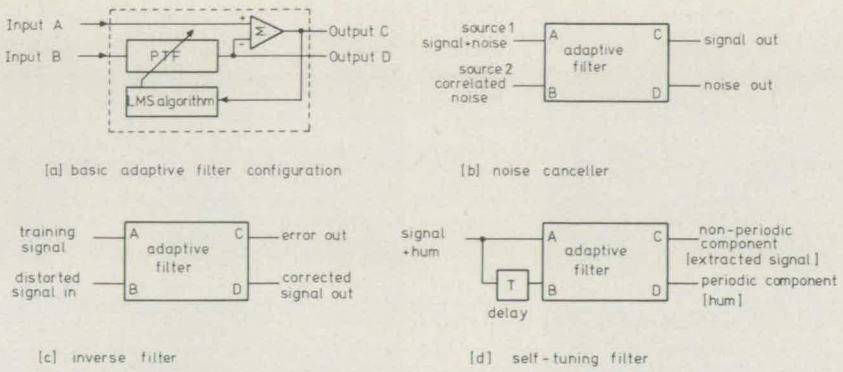


Fig. 13. Adaptive filter configurations in various application areas.

filter output; and Fig. 14(d) is the canceller output. The frequency amplitude spectrum of $d(t)$ is given in Fig. 14(e) with a corresponding spectrum for the error output in Fig. 14(f), which indicates cancellation of the first harmonic by 50 dB. The cancellation for the analogue update unit for the same test yields a value of about 25 dB: the poorer cancellation is attributed to the leaky nature of an analogue store produced with discrete components which results in excess error in the converged weight vector. A fully integrated analogue store should have improved characteristics and thus yield an improved cancellation result.

A general figure of merit for an adaptive filter is usually taken as the adaptivity,[24] χ, where:

$$\chi = 20 \log_{10} \left[\frac{\varepsilon(t)}{d(t)} \right] \qquad (9)$$

For the digital reference system described here χ was measured to be -50 dB.

Results for a simulated inverse filter and a self-tuning filter are now presented to demonstrate the versatility of an adaptive filter based on a p.t.f.

Figure 15 shows results obtained for a simulated inverse filter structure using the digital store version of the adaptive filter. The input signal $s(t)$ in this case was a linear f.m. signal swept from d.c. to the Nyquist frequency (shown in Fig. 15(a)). The training signal $d(t)$ was a single pulse having a width of one sample period, positioned at the end of the input signal sequence. The filter output after training is given in Fig. 15(b) showing good correspondence to the training signal. In this case the weight vector should match the time reverse of the input signal and the impulse response shown in Fig. 15(c) confirms this. A linear f.m. signal may be considered as a pulse with quadratic phase distortion, and is therefore a valid test signal to demonstrate inverse filter operation.

Tests carried out on the self-tuning filter showed that the system would tune accurately to a wide range of signals including monochromatic and wideband periodic signals. Figure 16 shows some typical results for the self-tuning filter where the input signal $s(t)$ (Fig. 16(a)) was a sinusoid contaminated by broadband noise. For the case shown, using a sinusoid at approximately 1 kHz, the broadband signal source was in fact a white noise generator which was limited in bandwidth between 1·5 kHz and 7 kHz. It is possible for the filter to converge when the noise spectrum is truly white from d.c. to Nyquist but this admits noise in the output in the region of the pass-band. Figure 16(b) shows the filter output D for this case and spectral analysis shows noise rejection to be between 25 and 30 dB after convergence. Figure 16(c) shows the cancellation output C where the sinusoidal component has been cancelled to a significant degree. The weight vector (shown in Fig. 16(d) is the expected matching sine-wave.

It should be noted that the weight vector shown in Fig. 16(d) is the signal which appears at the reference input of the p.t.f. and therefore contains all the gain and offset errors associated with the filter. Further, the impulse response cannot be perfectly predicted since the bandwidth of the input signals does not cover the entire system bandwidth; this means that a certain degree of latitude is available in the choice of weight values.

4 Conclusions

We have demonstrated in this paper how innovative c.c.d. and m.o.s. design techniques may be combined to realize an electronically-programmable transversal filter in compact form and having low power consumption. This filter configuration offers powerful signal processing capability and its programmable aspect allows flexible and fast control of the filter function, either remotely or internally, via a system-based microprocessor. The number of filter stages to be included on one chip must be a decision based upon the eventual application. For frequency filtering, relatively few filter points are required and must, in any case, be minimized to avoid dynamic range restrictions—possibly a 32- or 64-stage cascadable block is optimum. For matched filtering applications, however, many filter points may be necessary and it becomes desirable to realize as many points on one chip as possible.

In this paper, we have demonstrated the extremely

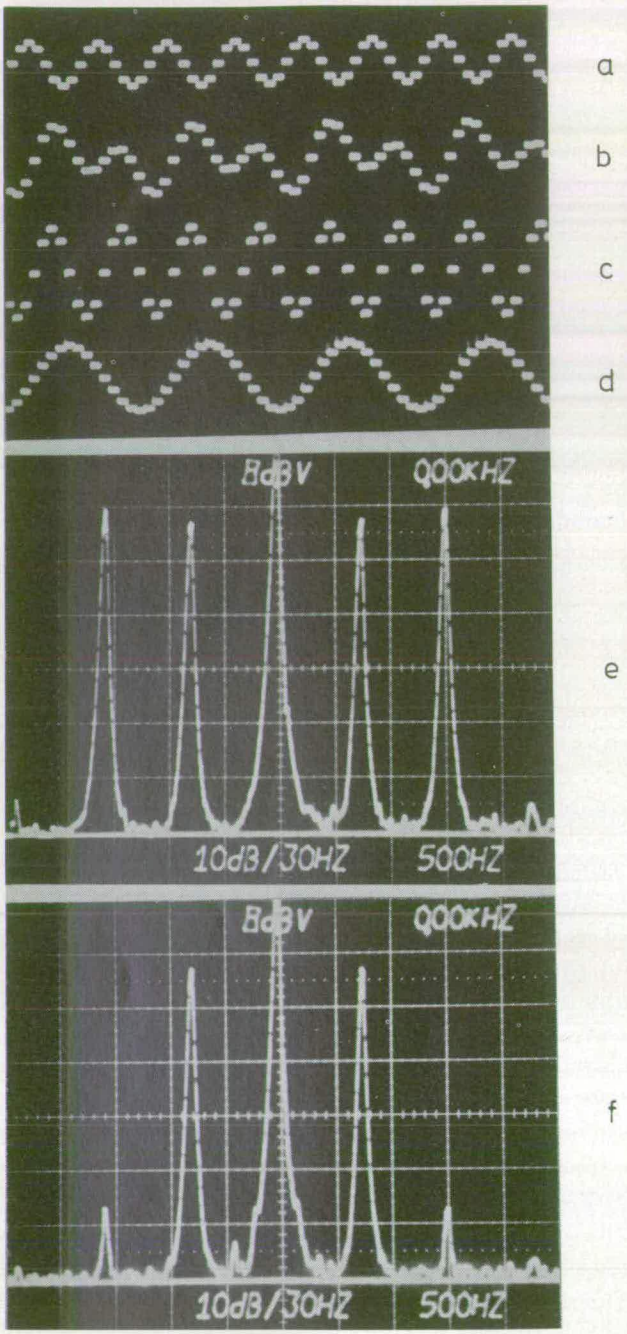


Fig. 14. Demonstration of the basic cancelling performance of the adaptive filter. (a) signal input $s(t)$. (b) target signal $d(t)$, desired sinusoid with added second harmonic distortion. (c) filter output, a phase and amplitude corrected version of the distortion. (d) canceller or error output. (e) spectrum of $d(t)$. (f) spectrum of canceller output.

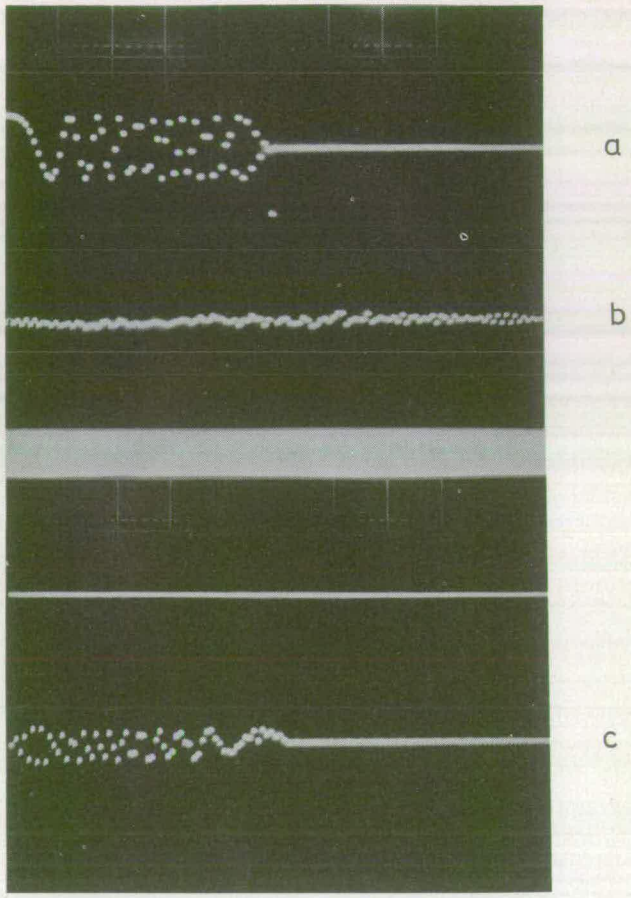


Fig. 15. Inverse filter operation. (a) input signal $s(t)$, dispersed pulse. (b) filter output, reconstituted pulse. (c) filter impulse response.

flexible performance of a prototype, 64-point electronically-programmable transversal filter for frequency- and matched-filtering applications. These results indicate that compact, low-power analogue c.c.d. p.t.f. sub-systems with TB products in excess of 1000 will be a reality in the near future, and will have an impact on such applications as mobile, high-resolution sonar signal processing. This is particularly evident in view of a 256-point p.t.f. chip already in design for high time-bandwidth applications. These filter structures are thus far only prototypes, but the 64-point device reported here and its contemporaries are proving that good performance parameters may be achieved in miniature form and that, with careful design, production devices must now be feasible.

We have also shown here that a 64-point adaptive filter may be realized using an analogue p.t.f. resulting in a compact unit of low-power dissipation (about 10 W for the complete system described). This type of processor could be used in a number of applications such as telephone line equalization, speech processing, cancellation of noise in transmitted speech, and a number of noise cancellation tasks in medical electronics. This is thought to be the first realistic demonstration of such a complex filtering system which could be integrated monolithically. Indeed, this work indicates that a dedicated adaptive filter with all necessary peripheral circuits, e.g. convergence multipliers and integrators, etc., could be produced in single chip form. However, such a dedicated design would not have the flexibility of a 'hybrid' unit as described here which could also be configured conveniently for other filtering functions such as spectrum analysis, matched filtering, etc. Also it would

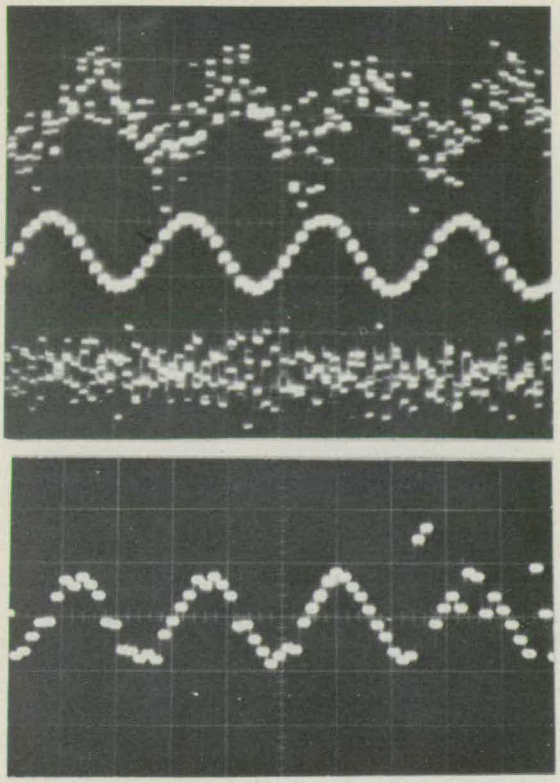


Fig. 16. Self-tuning filter performance. (a) sinusoid in bandlimited noise constituting the input signal. (b) filter output, noise rejection is approximately 25 dB. (c) error output showing cancellation of the periodic component. (d) converged weight vector.

not be able to achieve the same number of filter points because of the chip area needed for the peripheral circuitry.

5 Acknowledgements

This work has been carried out with the support of the Procurement Executive, Ministry of Defence: support from the Science Research Council is also acknowledged. Design and evaluation of the devices was undertaken in the Wolfson Microelectronic Institute, University of Edinburgh. Devices were processed by the Plessey Company. C. F. N. Cowan was supported by the Department of Education, Northern Ireland.

6 References

1 Boyle, W. S. and Smith, G. E., 'Charge-coupled semiconductor devices', *Bell Syst. Tech. J.*, **49,** pp. 587–93, April 1970.
2 Weste, N. and Mavor, J., 'M.o.s.t. amplifiers for performing peripheral integrated circuit functions', *IEE J. Electronic Circuits and Systems*, **1,** pp. 165–72, September 1977.
3 Hodges, D. A., Gray, P. R. and Brodersen, R. W., 'Potential of m.o.s. technologies for analog integrated circuits', *IEEE J. Solid-State Circuits*, **SC-13,** pp. 285–94, June 1978.
4 See, for example, RETICON TAD-32 Application Note, Reticon Corp., 910 Benicia Ave., Sunnyvale, California 94086.
5 Buss, D. D., Collins, D. R., Bailey, W. H. and Reeves, C. R., 'Transversal filtering using charge-transfer devices', *IEEE J. Solid-State Circuits*, **SC-8,** pp. 138–46, April 1973.
6 Denyer, P. B., Mavor, J. and Arthur, J. W., 'Miniature programmable transversal filter using c.c.d./m.o.s.t. technology', *Proc. IEEE*, **67,** pp. 42–50, January 1979.
7 Bosshart, P., 'An integrated analogue correlator using charge-coupled devices', Proc. of IEEE Int. Solid State Circuits Conf., 1976, pp. 198–9.
8 Haque, Y. A. and Copeland, M. A., 'Design and characterization of a real-time correlator', *IEEE J. Solid-State Circuits*, **SC-12,** pp. 642–9, December 1977.
9 Tiemann, J. J., Engeler, W. E. and Baertsch, R. D., 'A surface charge correlator', *IEEE J. Solid-State Circuits*, **SC-9,** pp. 403–9, December 1974.
10 Burke, B. E. and Lindley, W. T., 'New c.c.d. programmable transversal filter', *Electronics Letters*, **13,** pp. 521–3, September 1977.
11 Buss, D. D., Bailey, W. H. and Tasch, Jr., A. F., 'Signal processing applications of charge-coupled devices', Proceedings CCD 74, Edinburgh, pp. 179–97, September 1974.
12 Mavor, J., Arthur, J. W. and Denyer, P. B., 'Analogue c.c.d. correlator using monolithic m.o.s. multipliers', *Electronics Letters*, **13,** pp. 373–4, June 1977.
13 Denyer, P. B., Arthur, J. W. and Mavor, J., 'Monolithic programmable analogue c.c.d. transversal filter', *Electronics Letters*, **13,** pp. 373–4, November 1977.
14 Harp, J. G., Vanstone, G. F., MacLennan, D. J. and Mavor, J., 'Analogue correlators using charge coupled devices', Proceedings of CCD 75, pp. 229–35, September 1975.
15 Denyer, P. B. and Mavor, J., 'Design of c.c.d. delay lines with floating gate taps', *IEE J. Solid-State and Electron Devices*, **1,** pp. 121–9, July 1977.
16 Mavor, J. and Denyer, P. B., 'The design and development of c.c.d. programmable transversal filters and correlators', AGARD Conference Proceedings No. 230, pp. 2.8-1/15, October 1977.
17 Perkins, K. D. and Browne, V. A., 'Sub-micron gap metal gate technology for c.c.d.s', *Microelectronics J.*, **7,** pp. 14–22, January 1975.
18 Buss, D. D., Bailey, W. H., Holmes, J. D. and Hite, L. R., 'Charge transfer device transversal filters for communication systems', *Microelectronics J.*, **7,** pp. 46–53, January 1975.
19 Widrow, B. *et al.*, 'Adaptive noise cancelling: principles and applications', *Proc. IEEE*, **63,** pp. 1692–1716, December 1975.
20 Rabiner, L. R. and Gold B., 'Theory and application of digital signal processing' (Prentice-Hall, Englewood Cliffs, NJ, 1975).
21 Puckette, C. M., Butler, W. J. and Smith, D. A., 'Bucket-brigade transversal filters', *IEEE Trans. on Communications*, **22,** pp. 926–34, July 1974.
22 Cowan, C. F. N., Arthur, J. W. and Mavor, J., 'Noise cancellation and inverse filtering using a compact high-performance c.c.d. adaptive filter', *Electronics Letters*, **15,** pp. 35–7, January 1979.
23 Cowan, C. F. N., Mavor, J. and Arthur, J. W., 'Implementation of a 64-point adaptive filter using an analogue c.c.d. programmable filter', *Electronics Letters*, **14,** pp. 568–9, August 1978.
24 White, M. H., Mack, I. A., Borsuk, G. H., Lampe, D. R. and Kub, F., 'C.c.d. analogue adaptive signal processing', Proc. International Conference on 'The Technology and Applications of CCDs', San Diego, pp. 3A-1/14, October 1978.
25 Corl, D., 'A c.t.d. adaptive inverse filter', *Electronics Letters*, **14,** pp. 60–2, February 1978

Manuscript first received by the Institution on 25th June 1979 and in final form on 10th October 1979
(Paper No. 1931/CC 323)

UDC 621.382:535–1:621.383.8

Indexing Terms: Semiconductor devices, Integrated circuits, Charge-coupled devices, Pyroelectric detectors, Infra-red sensing

The applications of charge-coupled devices to infra-red image sensing systems

D. R. LAMB, Ph.D.*

and

N. A. FOSS, B.Sc., M.Sc.*

SUMMARY

The paper reviews the various ways in which c.c.d.s can be employed in i.r. sensing systems. These include: (i) monolithic structures fabricated using narrow band semiconductors such as HgCdTe or InSb, extrinsic silicon structures doped with deep-level impurities, and silicon Schottky barrier devices; (ii) hybrid structures in which the c.c.d. is used as the read-out mechanism from an array of, for example HgCdTe, PbTe, or pyroelectric detectors. The relative merits of these different approaches are compared and recent experiment results for many structures are quoted.

* *Honeywell Systems and Research Center, 2600 Ridgway Parkway, Minneapolis, Minn. 55413*

1 Introduction

Visible image sensors detect radiation reflected from objects of the scene and deal with situations where there is a low background flux and high scene contrast. Infra-red image sensors detect radiation emitted by objects in the scene and thus, in general terrestrial applications, have to detect very small temperature differences, frequently of the order of 0·1 K, against a background of around 300 K. As a result of this the two forms of imaging have developed in different ways.

In a visible imaging system, such as the silicon c.c.d. sensor, the scene is focused onto a large two-dimensional detector array having one detector per picture element. Small differences in responsivity of adjacent detectors are not important because the scene contrast is high. Also the photo-generated charge due to the low background photon flux is easily accommodated within the dynamic range of the c.c.d. for integration times compatible with standard television frame rates.

In i.r. imagers the detectors often have to be cooled to very low temperatures necessitating encapsulation of the sensor in a Dewar which leads to problems with the read-out interconnections from the i.r. detectors forming a thermal leakage path. This problem increases as the number of detectors is increased. Additionally, because of the low scene contrast, fixed-pattern noise in the detector array is a serious problem, (see Sect. 3) and in many cases some form of background suppression must be used because of the high background photon flux. For these reasons i.r. imaging is currently based on linear, or small two-dimensional arrays of detectors rather than large two dimensional arrays. To generate the number of picture elements for a complete picture field the image is mechanically raster-scanned over the detector array by means of rotating mirrors.

The initial concept of applying c.c.d.s to i.r. imaging systems arose from their ability to receive, store and manipulate charge packets to produce a multiplexed output. Thus one can envisage the use of a c.c.d. on the focal plane which will scan and read out the information from the detector array. In such a system a hybrid structure with separate detector and c.c.d. chips is used, thus taking advantage of both existing detector technology and the silicon integrated circuit technology.

With more ambitious programmes the aim is to produce monolithic i.r. c.c.d. imaging chips in which the photodetectors and c.c.d.s are fabricated in the same semiconductor material. Here the choice lies between producing silicon photodetectors by a technique which is compatible with the existing c.c.d. technology, or fabricating c.c.d.s in the narrow-band-gap semiconductors already used for i.r. detectors.

Both of these approaches are in fact being investigated; the former technique is, however, potentially of greater importance because it allows the possible use of silicon l.s.i. technology to make large two-

The Radio and Electronic Engineer, Vol. 50, No. 5, pp. 226–236, May 1980

0033–7722/80/050226+11 $1.50/0

MONOLITHIC 256-POINT PROGRAMMABLE TRANSVERSAL FILTER

Indexing terms: Charge-coupled-device circuits, Metal-oxide-semiconductor circuits, Transversal filters

A 256-point programmable transversal filter has been integrated and assembled for use within a 16-pin d.i.l. package. The filter is realised in *n*-channel c.c.d./m.o.s. technology and incorporates novel operational amplifiers and other peripheral circuitry on-chip.

Filter architecture: The integrated filter architecture shown in Fig. 1 is a direct realisation of the classic transversal filter. A

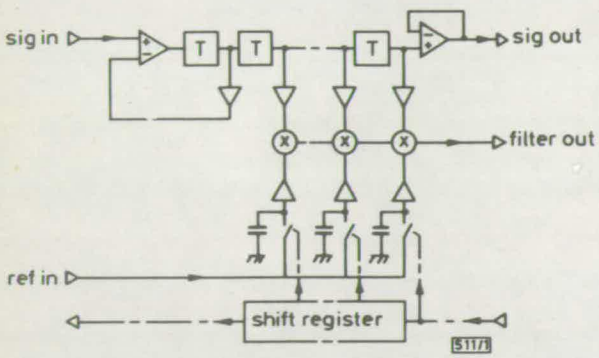


Fig. 1 *Filter architecture*

2-phase, double-level polysilicon-gate c.c.d. has been used as the signal delay element, with floating-gate-reset[1] taps incorporated at each stage on a pitch of 28 μm. The tap signals are buffered to drive the gates of single-transistor 4-quadrant m.o.s. multipliers.[2] The sources of these multipliers are commoned in a single current-summing bus which forms the filter output.

Analogue tap weights are applied to the multiplier transistor drains from an array of m.o.s. storage capacitors via buffer amplifiers. These reference storage capacitors are sequentially updated from a common reference input bus under the control of a dynamic digital shift register.

Operational Amplifiers: Two operational amplifiers have been included within the integrated circuit to enhance the operation of the c.c.d. One amplifier includes within a feedback loop the c.c.d. diode cutoff input stage and a dummy tap. It is used to linearise the signal waveform at this and subsequent identical taps,[3] and also ensures unity gain between signal input and tap outputs. The second operational amplifier is connected in voltage-follower mode at a final c.c.d. tap output to provide a buffered signal for directly cascading these devices.

The operational amplifiers, which are based on a clocked architecture,[4] are self-biasing and chopper-stabilised. Thus a linear lossless transfer function is achieved, automatically compensating for processing tolerances and parametric drift.

Topology: A photograph of the integrated circuit is shown in Fig. 2. It may be noted that the filter is split into two blocks of 128 points, which are mirror-imaged. Two 128-point c.c.d. registers are visible across the centre of the chip; these are connected by a single corner-turning diffusion at one end. The remainder of the cell circuitry, ending in the reference shift-register elements, is repeated on a pitch of 28 μm. The filter chip measures 4·0 mm × 3·8 mm (0·180 in × 0·150 in) and is assembled within a standard 16-pin d.i.l. package.

Peripheral requirements: Peripheral circuit requirements have been reduced to a minimum by incorporating as much necessary circuitry onto the chip as possible, thus making the filter easier to use. Essential external requirements are reduced to a 3-phase 15 V clock for the c.c.d. signal register, and a 2-phase 15 V clock for the reference shift register. These waveforms are conveniently generated using 2 standard c.m.o.s. logic packages. The filter output must be summed using a single common-base bipolar transistor stage. There is no other major external requirement.

Performance: Two photographs depicting the operation of the device are shown in Fig. 3. In each case the four traces show the signal input, the delayed signal output (for cascading), the programmed reference waveform and the filter output, respectively. Fig. 3*a* demonstrates full matched filter chirp compression over the Nyquist bandwidth at a clock rate of 100 kHz. Note the quality of both the delayed signal (for cascading) and the correlated output peak. Fig. 3*b* demonstrates square-wave correlation to produce a triangular output waveform again at 100 kHz clock rate.

The device performs uniformly over a range of signal clock rates from 500 Hz to 1 MHz. C.C.D. transfer inefficiency for

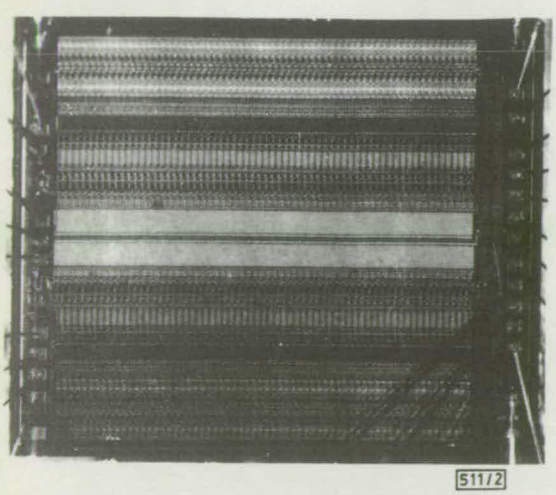


Fig. 2 *256-point chip photograph*

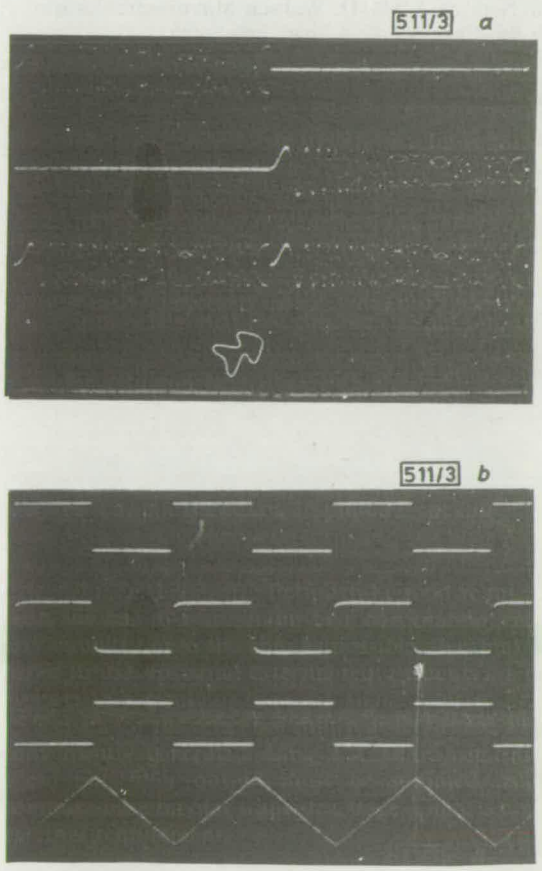


Fig. 3
a Matched filtering of chip waveforms T.B. = 128, f_c = 100 kHz
b Correlation of square waveforms f_c = 100 kHz

these surface-channel devices is 0·001 per tap and total harmonic distortion is −40 dB at the filter output, for signals of 1 V peak-peak. Power dissipation is below 300 mW.

Conclusions: The goal of a versatile easy-to-drive programmable transversal filter has been achieved through a combination of circuit engineering and modern process technology. This 16-pin 256-point filter offers considerable signal-processing power, and major applications are anticipated in all forms of matched-filtering tasks.

Acknowledgments: This work was sponsored by the Procurement Executive, MOD. Devices were processed by the Plessey Co. Ltd., Caswell.

P. B. DENYER*

Wolfson Microelectronics Institute
University of Edinburgh
Edinburgh, Scotland

* Now with Denyer Walmsley Microelectronics Ltd.

J. MAVOR

Department of Electrical Engineering
King's Buildings, University of Edinburgh
Edinburgh, Scotland

17th September 1979

References

1 DENYER, P. B., and MAVOR, J.: 'Design of c.c.d. delay lines with floating-gate taps', *IEE J. Solid-State & Electron. Devices*, 1977, **1**, pp. 121–129
2 DENYER, P. B., and MAVOR, J.: 'Miniature programmable transversal filter using CCD/MOS technology', *Proc. IEEE*, special issue on miniature filters, Jan. 1979, pp. 42–50
3 MACLENNAN, D. J., and MAVOR, J.: 'Novel technique for the linearisation of charge-coupled devices', *Electron. Lett.*, 1975, **11**, pp. 222–223.
4 DENYER, P. B., and MAVOR, J.: 'Novel m.o.s. differential amplifier for sampled data applications', *Electron. Lett.*, 1978, **14**, pp. 1–2
5 Application Note on WM2111, Wolfson Microelectronics Institute, King's Buildings, Mayfield Road, Edinburgh

0013-5194/79/220710-03$1.50/0